Special Issue: Proceedings of the Conference on
Surfaces and Interfaces in Mesoscopic Devices
(SIMD)'99
Guest Editor: J. P. Bird

Superlattices

and Microstructures
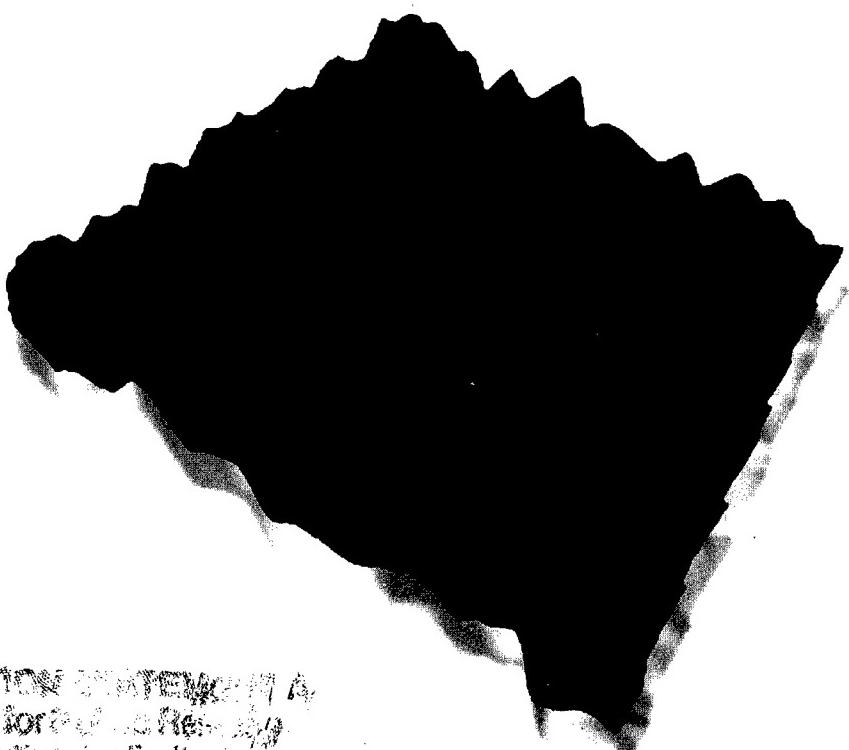
ACADEMIC PRESS

A Harcourt Science and Technology Company

Superlattices and Microstructures

An interdisciplinary Journal on the science and technology of nanostructures

NOTE: Manuscripts should be sent to the Editorial Office as described in the Instructions to Authors.

Superlattices and Microstructures is a journal devoted to the science and technology of synthetic microstructures, microdevices, surfaces and interfaces.

The last decade has seen rapid developments in the fabrication, characterization and conceptual understanding of synthetic microstructures in many different material systems including silicon, III-V and II-VI semiconductors, metals, ceramics and organics. The objective of this journal is to provide a common interdisciplinary platform for the publication of the latest research results on all such "nanostructures" with dimensions in the range of 1–100 nm; the unifying theme here being the dimensions of these artificial structures rather than the material system in which they are fabricated.

Cover picture shows a scanning tunneling microscope (STM) image of atoms on the surface of graphite—courtesy of Prof. R. Reifenberger's STM group in the Department of Physics at Purdue University, W. Lafayette, Indiana, U.S.A.

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT DATE	3. DATES COVERED *(From - To)*
June 2000	Final	Dec.1, 1999 – July 24, 2000

4. TITLE AND SUBTITLE

Superlattices and Microstructures in Mesosocopic Devices
Proceedings of SIMD 1999
Vol. 27, No. 5/6; May/June 2000

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-00-1-0059

5c. PROGRAM ELEMENT NUMBER
61153

6. AUTHOR(S)
Hess, Karl, Ferry, D. K.;
Bird, Jonathan

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Illinois
Beckman Institute
405 North Mathews Avenue
Urbana, Illinois 61801

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
Dr. L. R. Cooper, Code 312
Office of Naval Research
800 North Quincy
Arlington, VA 2217

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT

Approved for public release

13. SUPPLEMENTARY NOTES

14. ABSTRACT SIMD'99 (Surfaces and Interfaces of Mesoscopic Devices) has been held December 5-10, 1999 in Maui, Hawaii. Eleven Invited papers were given by D. D. Awschalom, University of California-Santa Barbara; S. DasSarma, University of Maryland; J. W. Lyding, University of Illinois; C. Quate, Stanford University, and others; 89 participants from 9 different countries attended. Topics of emphasis included electron spin coherence, entanglement and transport; metal insulator transitions in quantum dot arrays, phase braking times, Bohn trajectories and dissipative transport. A session was devoted to tunneling in extremely thin oxides and reliability questions for silicon technology.

Silicon-based nanoelectronics and nanomechanics were discussed in several papers including surface chemistry as resolved on the atomic scale by scanning tunneling microscopy.

Considerable progress was reported in all of these areas. All invited speakers attended and presented their topic and virtually all accepted contributed papers were given and are represented in the proceedings.

15. SUBJECT TERMS
Nanostrucutres, Interfaces, Mesoscopic Systems, Spintronics, Thin Oxides

16. SECURITY CLASSIFICATION OF:

a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
U	U	U	U U	

19a. NAME OF RESPONSIBLE PERSON
Karl Hess

19b. TELEPONE NUMBER *(Include area code)*
217-333-6362

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI-Std Z39-18

Guest Editorial

This special issue of Superlattices and Microstructures contains collected papers from the Third International Workshop on Surfaces and Interfaces in Mesoscopic Devices (SIMD'99), which was held on the Hawaiian island of Maui between December 6th and December 10th, 1999. Following on from the format of its predecessors, the focus of this workshop was on the fabrication of novel nanostructure devices, and the study of their transport properties. The workshop sessions included a number of key reviews and focused on spin transport in semiconductors, the 2D metal-insulator transition, reliability issue in MOSFETs, quantum-wire transport, single-electron-transistor-based quantum circuits, nanostructure fabrication throught STM-induced surface modification, and transport in carbon nanotubes. More than a hundred researchers from the United States, Japan, and Europe participated in the workshop. This represented an increase of roughly twenty five per cent over the corresponding figure for the previous workshop, which was held just two years earlier. Most encouragingly, this higher level of participation was achieved through increased attendance of researchers from around the world, helping to make this workshop one of truly international significance. It is therefore my hope that the content of these proceedings will serve as a valuable, and timely, reference for researchers in this field.

On behalf of the organizing committee of the workshop, I would like to take this opportunity to acknowledge the kind of sponsorship of the workshop by the Office of Naval Research and the Beckman Institute of the University of Illinois. As Guest Editor, I would also like to personally thank Cheryl Ream (Arizona State University) and Sara Starkley (University of Illinios) for their very kind assistance in helping to oversee the publication of these proceedings.

Jonathan Bird
Department of Electrical Engineering
Arizona State University

Spintronics: electron spin coherence, entanglement, and transport

S. Das Sarma, Jaroslav Fabian, Xuedong Hu, Igor Žutić

Department of Physics, University of Maryland, College Park, MD 20742-4111, U.S.A.

The prospect of building spintronic devices in which electron spins store and transport information has attracted strong attention in recent years. Here we present some of our representative theoretical results on three fundamental aspects of spintronics: spin coherence, spin entanglement, and spin transport. In particular, we discuss our detailed quantitative theory for spin relaxation and coherence in electronic materials, resolving in the process a long-standing puzzle of why spin relaxation is extremely fast in Al (compared with other simple metals). In the study of spin entanglement, we consider two electrons in a coupled GaAs double-quantum-dot structure and explore the Hilbert space of the double dot. The specific goal is to critically assess the quantitative aspects of the proposed spin-based quantum dot quantum computer architecture. Finally, we discuss our theory of spin-polarized transport across a semiconductor/metal interface. In particular, we study Andreev reflection, which enables us to quantify the degree of carrier spin polarization and the strength of interfacial scattering.

© 2000 Academic Press

Key words: spintronics, spin coherence, spin relaxation, spin-hot-spot model, spin entanglement, electron exchange, spin transport, Andreev reflection, spin tunneling.

1. Introduction

There has been a great deal of recent interest in the concept of spintronics [1] where active control and manipulation of electron spin in semiconductors and metals provide the basis of a novel quantum technology. Possible applications of spintronics include high-speed magnetic filters, sensors, quantum transistors, and spin qubits for quantum computers [2,3]. More fundamental research will, however, be needed before practical spintronic devices can be demonstrated, as much remains to be understood about spin coherence, spin dynamics, and spin transport. In this paper we discuss some of our recent theoretical work on understanding spin dynamics in electronic materials.

The existing spintronic architectures [1,4] and the proposed solid-state quantum computing schemes [3] rely on the relatively long spin coherence times of conduction electrons. Indeed, in the simplest spintronic scheme—the spin injection [5]—electrons with a definite spin polarization are supplied into a nonmagnetic metal or semiconductor from a ferromagnetic electrode. The farther (longer) the electrons in the nonmagnetic sample carry the spin coherence, the more useful the device is. Similarly, if an electron spin represents a qubit in a solid-state quantum computer, the longer the spin survives, the more reliably it can store information. The question of how spins of mobile electrons (and holes) lose their spin coherence is thus of the utmost importance for spintronic technology and for solid-state quantum computing. Unfortunately, the physical

picture of spin decoherence (or relaxation) we now have is far from complete. Most information comes from experiments, but experimental data are still scarce and often incomplete. The existing theories seem to provide a broad conceptual framework for understanding spin decoherence in metals and semiconductors, but the acute absence of realistic calculations for concrete materials makes it difficult to validate these theories. The hope is that with more complete experimental and theoretical understanding we will be able to choose or build materials with the longest decoherence times possible. In this paper we discuss, as an example, our detailed quantitative theory [6–8] of spin relaxation and coherence in a simple metal, Al.

One challenge of spintronics is to study the possibility of using (electron or nuclear) spins as quantum bits (qubits) in a quantum computer (QC). QC has drawn growing attention in recent years because it can deliver significant speed-up over classical computers [2] for some problems due to inherent superposition and entanglement of a quantum system. Various QC architectures have been proposed, including several solid state models that may possess the important feature of scalability. We study a QC model [3] in which a double quantum dot in the GaAs conduction band serves as the basic elementary gate for a QC with the electron spins in the dots as qubits. The two-electron exchange coupling provides the necessary two-qubit entanglement required for quantum computation. Using a molecular orbital approach [9], we determine the excitation spectrum of two horizontally coupled quantum dots with two confined electrons, and study its dependence on an external magnetic field [10]. We particularly focus on the electron exchange coupling and double occupation probability, which are two crucial parameters for a QC architecture.

Since potential spintronic devices are typically heterojunctions [5], it is important to understand how transport across the interface between different materials depends on the degree of carrier spin polarization and the interfacial transparency. Significant progress has been made in manipulating spin dynamics in semiconductors [11], including various methods to create spin-polarized carriers [12], such as employing a novel class of ferromagnetic semiconductors [13] and spin injection from a ferromagnet [14]. These advances, together with tunable electronic properties (such as carrier density and Fermi velocity) and well-established fabrication techniques, provide a compelling reason to study hybrid semiconductor structures in the context of spintronics [1]. In particular, we consider spin-polarized transport in the semiconductor/superconductor hybrid structures, which for low applied bias is governed by Andreev reflection [15]. Our aim is also to motivate study of the interplay between spin polarization and Andreev reflection in other areas, such as mesoscopic physics and quantum computing.

2. Spin coherence in electronic materials

Spins of conduction electrons decay because of the spin–orbit interaction and momentum scattering. At low temperatures ($T \lesssim 20$ K) spin relaxation is caused by impurity scattering and is temperature independent. At higher temperatures electrons lose spin coherence by colliding with phonons (phonons can induce a spin flip because in the presence of a spin–orbit coupling electronic Bloch states are not spin eigenstates). Spin relaxation rate $1/T_1$ increases as temperature increases, with the growth becoming linear above the Debye temperature. This mechanism, discovered by Elliott [16] and Yafet [17], is the most important spin relaxation mechanism in metals and semiconductors with inversion symmetry. It gives typical values of T_1 on the nanosecond scale, in agreement with experiment. To our knowledge the longest T_1 in a metal has been reported to be a microsecond, in a very pure Na sample at low temperatures [18]. The situation is much more complicated in semiconductors [8]. Many interesting semiconductors such as GaAs lack inversion symmetry, so other mechanisms, in addition to the Elliott–Yafet one, become important [8]. These mechanisms operate differently in different temperature regions, doping, and magnetic, strain, and confinement fields, so that sorting out the relevant mechanism(s) for a given material is a tremendous task which is yet to be carried out. Magnitudes of T_1 in semiconductors are also typically nanoseconds, but recent experimental studies [11] in II–VI and III–V systems show that T_1 can be artificially enhanced [8]. It seems, however, that intrinsic elec-
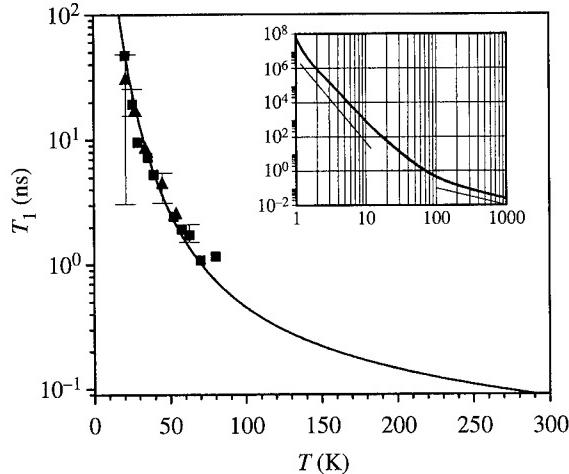
Fig. 1. Calculated phonon-induced spin relaxation time T_1 in Al as a function of temperature T. Symbols with error bars are experimental data from [5, 19] and the inset is a larger-scale log–log plot. There is good agreement between the theory and the experiment, but the absence of experimental data at temperatures above 100 K makes the calculation a prediction which is particularly useful for spintronic applications at room temperature.

tronic properties may not allow T_1 in technologically interesting materials to be longer than a microsecond at room temperature.

Spin relaxation is very sensitive to the electronic band structure. Our *ab initio* calculation [6, 7] of T_1 in Al, whose result is shown in Fig. 1, shows that bandstructure anomalies such as the Fermi surface crossing of a Brillouin zone boundary or an accidental degeneracy line, can enhance spin relaxation (reduce T_1) by orders of magnitude. Since such anomalies are ubiquitous in polyvalent metals (the Fermi surfaces of monovalent metals are well contained within the first Brillouin zone boundary and thus are free of anomalies), we gave them a special name: spin hot spots. Whenever an electron jumps in or out (as a result of a collision) of a spin hot spot, the electron's chance of flipping spin is greatly enhanced. As a result, $1/T_1$ in Al and other polyvalent metals is much greater than what one would naively expect. This is indeed what is measured [20]: While monovalent alkali and noble metals have their spin relaxation rates in accordance with simple estimates based on the Elliott–Yafet theory, polyvalent metals (only Al, Be, Mg, and Pd have been measured so far) have $1/T_1$ larger than expected by typically one to three orders of magnitude. Take as an example Al and Na. They have similar atomic numbers, so one would expect that their corresponding spin–orbit couplings would also be similar (as is the case in the atomic state [17]), giving similar spin relaxation rates. However, the corresponding T_1 at the Debye temperatures (150 K for Na and 390 K for Al) are about 20 ns in Na and 0.1 ns in Al! This huge difference is caused by the presence (absence) of spin hot spots in polyvalent Al (monovalent Na).

The current fashion for electron spin aside, spin relaxation in electronic materials is a beautiful and important subject of its own. The field itself began in the 1950s with the advent of CESR (conduction electron spin resonance), but after some initial breakthroughs the subject went dormant until the current surge inspired by spintronics. The experimental focus has so far been on the simplest elemental metals like the alkali and the noble metals, and on just a handful of interesting semiconductors. This is understandable from the point of view of technological applications, but not quite right from the point of view of fundamental physical understanding. What is clearly needed is catalogue of temperature-dependent spin relaxation times for different metals and semiconductors, a systematic study of the effects of impurities, alloying, and surfaces and interfaces. Theory should have the same goal: performing realistic calculations of T_1 for different metals within the existing framework laid out by Elliott and Yafet, and similarly for semiconductors taking into account other spin relaxation mechanisms as well.
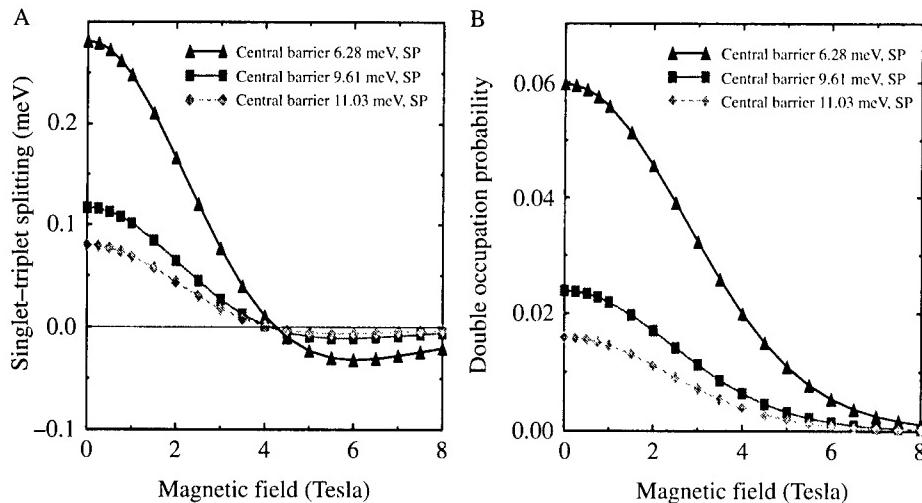
Fig. 2. Calculated magnetic field dependence of A, the spin singlet–triplet splitting and B, double occupation probability in the singlet state for a two-electron double quantum dot. The dot size is 30 nm in radius, and the inter-dot distance 40 nm. There is a single–triplet crossing at about 4 Tesla, which is a result of competing Coulomb repulsion, exchange, and single particle kinetic and potential energies. DOP is relatively high at low magnetic fields, but is quickly suppressed as magnetic field increases.

3. Electron entanglement through exchange interaction in a double quantum dot

The exchange coupling (the splitting of the lowest singlet and triplet states) and the double occupation probability (the probability that the two electrons occupy the same orbital state in one dot) in a double dot are two important parameters for the spin-based quantum dot quantum computer (QDQC). The exchange coupling between two electrons establishes the necessary entanglement between spins, and determines how fast quantum gates can be. Quantum computation has very low tolerance for errors (it requires an error rate below 10^{-4}), so that precise control and small errors are imperative for a QC. In a QDQC, individual quantum dots are tags which distinguish different qubits. If during a gating action two electrons jump onto a single quantum dot, their original tag information will be lost, which will result in an error. Thus, in designing a QDQC, double occupation probability (DOP) has to be minimized for the states that belong to the QDQC Hilbert space. Figure 2 shows our numerical results on the magnetic field dependence of (A) the exchange coupling with three different central barrier heights, and (B) the ground state DOP. The latter clearly decreases as **B** field increases. Physically, as **B** increases, the single-electron atomic wavefunctions are squeezed so that the inter-dot wavefunction overlap decreases, while the 'on-site' Coulomb repulsion energy for a single dot increases. The ground state DOP can also be seen in Fig. 2 to decrease significantly with increasing central barrier strength separating the two dots, as one would expect.

As shown in Fig. 2, at zero magnetic field it is difficult to have both a vanishing DOP for a small error rate and a sizeable exchange coupling for fast gating, because the exchange coupling and the DOP have similar dependence on the inter-dot barrier and inter-dot distance. On the other hand, finite magnetic fields may provide finite exchange coupling for QC operations with small errors. However, a finite magnetic field will produce a Zeeman splitting in the triplet state, causing additional phase shifts. Therefore, a swap gate [3] in a QC would have to include additional single-qubit operations to correct the effects of these phase shifts [10]. This added complexity inevitably prolongs the gating time of a two-qubit operation, which in turn increases the chance of an error due to spin relaxation. Another implication of a finite magnetic field is small exchange coupling—about an order of magnitude smaller than that at zero field. This means that the two-qubit operations will last as long as 10 ns, requiring the spin coherence time to be longer than 10 μs in a semi-

conductor quantum dot. Whether GaAs or other electronic materials can provide such favorable environment for electron spins is yet to be determined, and many other questions need to be answered before practical spin-based QDQC can be realized. These questions include, but are not limited to, the effects of stray fields, the implication of a chosen geometry, and the effects of external noise introduced through active control.

4. Spin transport

The presence of spin-polarized carriers gives rise to both modified charge transport and intrinsic spin transport, absent in the unpolarized case. Each of these aspects provides information about the degree of spin polarization which can be utilized in spintronics. Here, we focus on the transport of spin-polarized carriers across the semiconductor/metal interface where the metal is in the superconducting state. The study of semiconductor/superconductor (Sm/S) hybrid structures has several important ramifications. Already in the context of spin-unpolarized transport [21], it has been demonstrated [22] that this configuration can be used to examine the interfacial transparency which for a Sm/normal metal is typically limited by a native Schottky barrier. In the presence of spin-polarized carriers, Sm/S structure can also serve to quantify the degree of spin polarization of a semiconductor and probe both potential and spin-flip interfacial scattering [12]. To understand such sensitivity to spin polarization and different types of interfacial scattering it is important to consider the process of Andreev reflection [15] which governs the low bias transport. In this two-particle process, an incident electron of spin $\sigma =\uparrow, \downarrow$ on a Sm/S interface is reflected as a hole belonging to the opposite spin subband, back to the Sm region while a Cooper pair is transferred to the superconductor. The probability for Andreev reflection at low bias voltage is thus related to the square of the normal state transmission coefficient and can have stronger dependence on the junction transparency than the ordinary single-particle tunneling. For spin-polarized carriers, with different populations in two spin subbands, only a fraction of the incident electrons from a majority subband will have a minority subband partner in order to be Andreev reflected. In the superconducting state, for an applied voltage smaller than the superconducting gap, single particle tunneling is not allowed in the S region and the modification of the Andreev reflection amplitude by spin polarization or junction transparency will be manifested in transport measurements.

Prior to work from Ref. [12] the spin-dependent Andreev reflection was addressed in the context of ferromagnet/S junctions [23] and calculations were performed assuming the equality of the effective masses [24] in the two regions across the interface. Such an assumption is inadequate for the Sm/S hybrid structures. We adopt here the scattering approach from Ref. [12] and solve the Bogoliubov–de Gennes [12, 24] equations in a ballistic regime. At the flat interface between the Sm and S region we model the interfacial scattering by Z_σ and F, orbital and spin–flip scattering strengths, respectively. $Z_\uparrow \neq Z_\downarrow$ can describe magnetically active interface and the effects of spin-filtering. We represent spin polarization by X, the ratio of spin subband splitting and the Fermi energy. In Fig. 3 we give normalized low-temperature results [12] for the three-dimensional charge conductance, $G_3 \equiv G_{3\uparrow} + G_{3\downarrow}$, as a function of the ratio of bias voltage, eV, and the superconducting gap, Δ. Displayed charge conductance, which is calculated for vanishing interfacial scattering strength, depends strongly on the spin polarization X. The inset shows the effect of orbital and spin-flip scattering at a fixed spin polarization of $X = 0.4$ [12], $G_{S3} \equiv G_{S3\uparrow} - G_{S3\downarrow}$. To study the intrinsic spin transport, it is convenient to define spin conductance. For $eV < \Delta$ and any spin polarization $G_{S3} = 0$, since $G_{S3\uparrow,\downarrow}$ are each proportional to the corresponding spin component of the quasiparticle current and that there is no quasiparticle tunneling below the superconducting gap. For the unpolarized case, $X = 0$, $G_{S3\uparrow} \equiv G_{S3\downarrow}$ and the spin conductance vanishes identically. For $eV > \Delta$, G_{S3} is a sensitive function of X and could be used to determine the degree of the spin polarization. Experimental studies of spin-polarized Sm/S junctions should provide an important test for feasibility of spintronic devices based on hybrid semiconductor structures, as well as stimulate future theoretical studies considering, for example, nonequilibrium processes, realistic band structure, and diffusive regime.
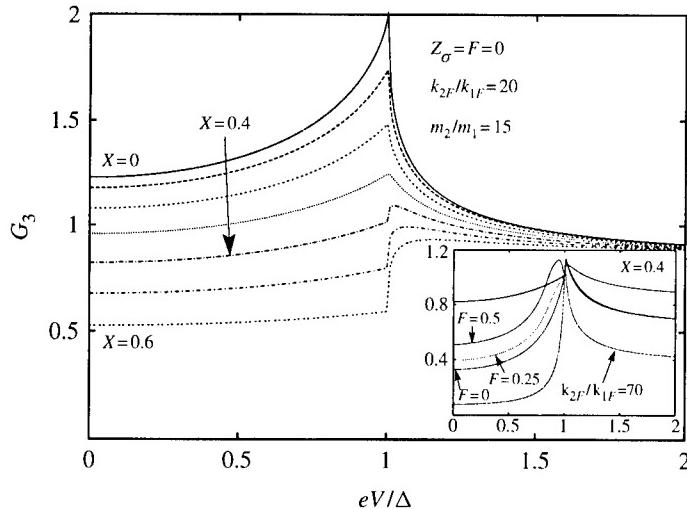
Fig. 3. Normalized charge conductance $G_3(eV/\Delta)$. Curves from top to bottom represent $X = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6$ at $Z_\sigma = F = 0$. The inset shows $X = 0.4$ results. The upper four curves (from top to bottom at zero bias) have Z_σ and F values of (0,0), (0.5,0.5), (0.5,0.25), and (0.5,0). For the bottom curve, which corresponds to $k_{2F}/k_{1F} = 70$, the Z_σ and F values are (0,0). Here the effective masses and the Fermi wave vectors are denoted by m_i and k_{if}, where $i = 1, 2$ correspond to the Sm and S regions, respectively.

5. Conclusion

We studied several issues related to spintronics. First, we demonstrated the importance of band structure effects in spin relaxation and established that special subtle features (spin hot spots) of electronic structure have profound effects on the magnitude of the spin relaxation rate in electronic materials. Since spin hot spots can be artificially induced, our work also shows a way of tailoring spin dynamics of conduction electrons. Next, our study of a quantum dot hydrogen molecule showed that the goal of having both a reasonable exchange coupling and a vanishingly small error rate can only be achieved at finite magnetic fields (4–8 Tesla), and one has to consider many factors (such as fast gating time, precise control, low error rate, etc) to produce a realistic spin-based quantum computer. Finally, we demonstrated that the low-temperature spin-polarized transport in Sm/S structures may serve as a sensitive and quantitative probe for determining the degree of spin polarization and the strength of interfacial scattering. In contrast to the unpolarized case, the junction transparency can be enhanced with the increase of the Fermi velocity mismatch in the two regions.

References

[1] G. Prinz, Phys. Today **48**, 58 (1995); Science **282**, 1660 (1998).

[2] A. Steane, Rep. Prog. Phys. **61**, 117 (1998).

[3] D. Loss and D. DiVincenzo, Phys. Rev. **A57**, 120 (1998); G. Burkard, D. Loss, and D. DiVincenzo, Phys. Rev. **B59**, 2070 (1999).

[4] J. Gregg *et al.*, J. Magn. Magn. Mater. **175**, 1 (1997).

[5] M. Johnson and R. H. Silsbee, Phys. Rev. Lett. **55**, 1790 (1985).

[6] J. Fabian and S. Das Sarma, Phys. Rev. Lett. **81**, 5624 (1998); J. Appl. Phys. **85**, 5057 (1999).

[7] J. Fabian and S. Das Sarma, Phys. Rev. Lett. **83**, 1211 (1999).

[8] J. Fabian and S. Das Sarma, J. Vac. Sci. Technol. **B17**, 1708 (1999).

[9] R. McWeeny, *Methods of Molecular Quantum Mechanics* (Academic Press, San Diego, 1992).

[10] X. Hu and S. Das Sarma, LANL Preprint quant-ph/9911080. To appear in Phys. Rev. **A**.

[11] D. D. Awschalom and J. M. Kikkawa, Phys. Today **52**, 33 (1999).

[12] I. Žutić and S. Das Sarma, Phys. Rev. **B60**, 16322 (1999).

[13] H. Ohno, Science **281**, 951 (1998).

[14] P. R. Hammar, B. R. Bennett, M. J. Yang, and M. Johnson, Phys. Rev. Lett. **83**, 203 (1999).

[15] A. F. Andreev, Zh. Eksp. Teor. Fiz. **46**, 1823 (1964).

[16] R. J. Elliott, Phys. Rev. **96**, 266 (1954).

[17] Y. Yafet, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1963) Vol. 14.

[18] W. Kolbe, Phys. Rev. **B3**, 320 (1971).

[19] D. Lubzens and S. Schultz, Phys. Rev. Lett. **36**, 1104 (1976).

[20] P. Monod and F. Beuneu, Phys. Rev. **B19**, 911 (1979).

[21] C. J. Lambert and R. Raimondi, J. Phys.: Condens. Matter **10**, 901 (1998); C. W. J. Beenakker, Rev. Mod. Phys. **69**, 731 (1997), give reviews of unpolarized transport in Sm/S junctions and address Andreev reflection in mesoscopic physics.

[22] S. De Franceschi, F. Giazotto, F. Beltram, L. Sorba, M. Lazzarino, and A. Franciosi, Appl. Phys. Lett. **73**, 3890 (1998).

[23] V. A. Vas'ko *et al.*, Appl. Phys. Lett. **73**, 844 (1998); R. J. Soulen Jr. *et al.*, Science **282**, 85 (1998); S. K. Upadhyay *et al.*, Phys. Rev. Lett. **81**, 3247 (1998).

[24] M. J. M. de Jong and C. W. J. Beenakker, Phys. Rev. Lett. **74**, 1657 (1995); J.-X. Zhu *et al.*, Phys. Rev. B **59**, 9558 (1999); I. Žutić and O. T Valls, Phys. Rev. **B60**, 6320 (1999); *ibid*, **61**, 1555 (2000); S. Kashiwaya *et al.*, Phys. Rev. **B60**, 3572 (1999).

Demonstration of electrical spin injection into a semiconductor using a semimagnetic spin aligner

G. Schmidt, R. Fiederling, M. Keim, G. Reuscher, T. Gruber, W. Ossau,
A. Waag, L. W. Molenkamp

Physikalisches Institut, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany

Spin injection into semiconductors has been a field of growing interest during recent years, because of the large possibilities in basic physics and for device applications that a controlled manipulation of the electrons spin would enable. However, it has proven very difficult to realize such a spin injector experimentally. Here we demonstrate electrical spin injection and detection in a GaAs/AlGaAs p-i-n diode using a semimagnetic II–VI semiconductor ($Zn_{1-x-y}Be_xMn_ySe$) as a spin aligner. The degree of circular polarization of the electroluminescence from the diode is related to the spin polarization of the conduction electrons. Thus, it may be used as a detector for injected spin-polarized carriers. Our experimental results indicate a spin polarization of the injected electrons of up to 90% and are reproduced for several samples. The degree of optical polarization depends strongly on the Mn concentration and the thickness of the spin aligner. Electroluminescence from a reference sample without spin aligner as well as photoluminescence after unpolarized excitation in the spin aligner sample show only the intrinsic polarization in an external magnetic field due to the GaAs bandstructure. We can thus exclude side effects from Faraday effect or magnetic circular dichroism in the semimagnetic layer as the origin of the observed circularly polarized electroluminescence.

© 2000 Academic Press

Key words: spin injection, dilute magnetic semiconductors.

Introduction

Recently magnetoelectronics and spin-polarized transport have developed very rapidly [1]. However, while giant magneto resistance (GMR) in all-metal systems is already commercialized and tunneling magnetic random access memory (TRAM) is close to the market, spin injection into semiconductors has seen no considerable progress. Although various efforts have been made, all experiments published up to now show only very small results [2, 3], which are difficult to interpret and may as well contributed to magnetoresistance or Hall effect. These experiments were mainly done using a simple device geometry with two ferromagnetic metal contacts on a semiconductor two-dimensional electron gas (2DEG) where the device resistance is expected to depend on the relative magnetization [3, 4] of the contacts.

Moreover, we have recently shown theoretically, that in the range of diffusive transport and linear response, these spin injection experiments will yield no detectable effects if typical metallic ferromagnets like Ni, Fe, or Co are used as contact materials [5], because of the finite degree of the spin polarization of the current
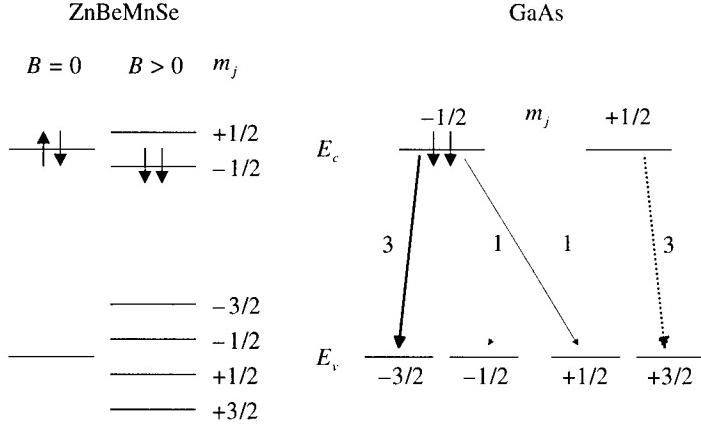
Fig. 1. Zeeman splitting in ZnBeMnSe in an external magnetic field (left) and possible radiative transitions in GaAs (right). In the ZnBeMnSe the electrons relax into the lower Zeeman level. In the GaAs the recombination probability with heavy holes is three times as high as for the light hole transition resulting in circularly polarized emission if the electrons are spin polarized.

carrying electrons in these materials. We have therefore followed a different experimental approach, and used a semimagnetic semiconductor as a spin injection contact instead of a ferromagnetic metal. The current was injected into a GaAs/AlGaAs light-emitting diode, where the spin polarization could be detected via the circular optical polarization of the emitted light [6]. In the following, the detection principle is explained. The device as well as its fabrication process is described and the experimental results are presented.

1. The device

In GaAs, two radiative recombination mechanisms for holes and electrons are possible. An electron from the nondegenerate conduction band ($m_j = \pm 1/2$) may recombine either with a light hole ($m_j = \pm 1/2$) or a heavy hole ($m_j = \pm 3/2$), with the probability of the heavy hole process being three times as high as for the light hole process. The transitions are spin selective, e.g. a spin-up electron ($m_j = +1/2$) can either recombine with a spin-up heavy hole ($m_j = +3/2$) resulting in a σ^+ or with a spin-down ($m_j = -1/2$) light hole resulting in a σ^- circularly polarized photon and vice versa (Fig. 1). If the electrons in the conduction band are partially spin polarized the spin polarization is defined as follows:

$$\alpha = \frac{n^\uparrow - n^\downarrow}{n^\uparrow + n^\downarrow} \tag{1}$$

where $n_\uparrow$ ($n_\downarrow$) is the number of the current carrying electrons with spin-up (spin-down), respectively. Radiative recombination processes with unpolarized holes will lead to the following circular polarization of the emitted radiation:

$$P_{circ}^{opt} = \frac{\sigma^+ - \sigma^-}{\sigma^+ + \sigma^-} = \frac{(3n^\uparrow + n^\downarrow) - (3n^\downarrow + n^\uparrow)}{(3n^\uparrow + n^\downarrow) + (3n^\downarrow + n^\uparrow)}. \tag{2}$$

For degenerate light and heavy hole states the optical circular polarization will thus be half the value of the spin polarization of the recombining electrons ($\alpha = 2P_{circ}^{opt}$). If the degeneracy is lifted and we see only the heavy hole transition the degree of optical polarization and of spin polarization are equal ($\alpha = P_{circ}^{opt}$). If a GaAs LED is driven by spin-polarized electrons from the n-side, while unpolarized holes are supplied from the p-side, the spin polarization may thus be determined from the degree of circular polarization P_{circ}^{opt} of the electroluminescence signal. ZnMnBeSe is a semimagnetic II–VI semiconductor which exhibits, at low
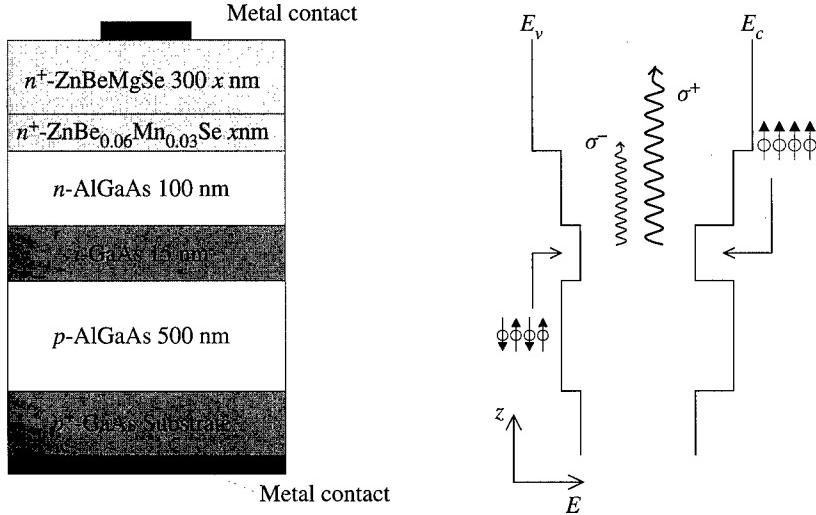
Fig. 2. Layer sequence (left) and band structure (right) of the spin aligner LED. The device consists of a GaAs/AlGaAs LED which has a spin aligning n-contact. The electrons are aligned in the II–VI layer and recombine in the GaAs with unpolarized holes coming from the GaAs substrate.

temperature, a so-called giant Zeeman splitting in the valence band and the conduction band if an external magnetic field is applied (Fig. 1). This large splitting will lead to the relaxation of *all* conduction electrons into the lower Zeeman level while passing through a ZnMnBeSe layer. Thus, in contrast wit the ferromagnetic metals commonly used as spin injector, our II–VI spin aligner injects a 100% spin-polarized current into the nonmagnetic semiconductor layer. The material can thus be considered as a good candidate for a spin aligner.

We have fabricated several LEDs. All structures consisted of a GaAs/AlGaAs p-i-n diode grown on a p-doped GaAs substrate. On top of the diode, a highly n-doped $Zn_{0.91}Be_{0.06}Mn_{0.03}Se$ spin aligner was grown (Fig. 2). Depending on the thickness of the spin aligner d_{sa} an additional n-doped $Zn_{1-x-y}Be_xMg_ySe$ was added, to obtain an overall thickness of the II–VI layer of 300 nm. This was done in order to achieve identical current spreading in the different devices. The semiconductor layers were grown in in two different MBE chambers for the II–VI and the III–V layers, respectively, and transferred in vacuum. To complete the devices, the structures were covered by a 100 nm Aluminum layer deposited *in situ*, again without breaking the UHV. After removing the samples from the MBE system, a Mesa was etched and central Al contact pads were fabricated using thermal evaporation, optical lithography and lift-off. The LEDs fabricated this way were bonded and characterized in a He[4] bath cryostat.

2. Experimental results

Three LEDs showed distinct circular polarization in the electroluminescence signal (Fig. 3). For a spin aligner thickness d_{sa} of 100 and 300 nm, P_{circ}^{opt} was about 45%. For a thickness of 3 nm P_{circ}^{opt} is still more than 10% while for 300 nm of nonmagnetic $Zn_{1-x-y}Be_xMg_ySe$ it was absent within the detection limit (Fig. 1). We attribute the large degree of optical polarization with a spin aligner contact to spin injection. The increasing polarization with the spin aligner thickness and a saturation at higher magnetic fields are well in agreement with theory, as the spin aligner exhibits a finite spin relaxation length and the spin polarization will be 100% as soon as the b-field dependent Zeeman splitting is large compared with kT. As the GaAs layer was designed such that degenerate light and heavy hole states can be expected this would indicate a spin polarization of up to 90%. The Faraday effect and giant magnetic circular dichroism [7] were ruled out
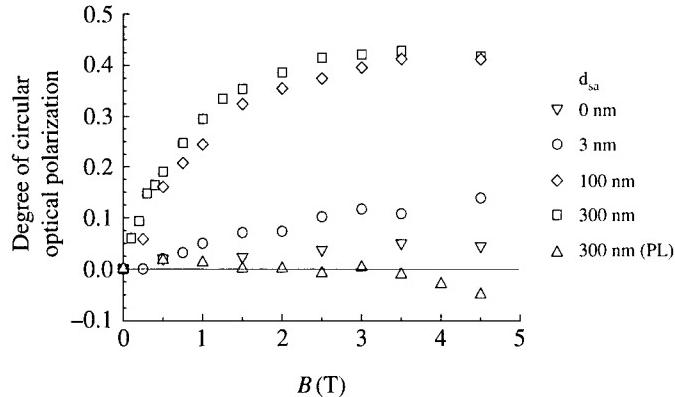
Fig. 3. The degree of circular optical polarization is plotted over the external magnetic field for LEDs with different spin aligners. The electroluminescence signal shows a small polarization for $d_{sa} = 0$ (no spin aligner). P_{circ}^{opt} increases with the spin aligner thickness and saturates for $d_{sa} \approx 100$ nm. The PL signal for $d_{sa} = 300$ nm shows small polarization with opposite handedness originating from the Zeeman splitting.

by an additional experiment in which unpolarized electrons and holes where excited in the well by optical excitation. In a magnetic field, the photoluminescence signal showed only a small circular polarization with a handedness opposite to the polarization which is caused by spin injection. The origin of this polarization is the intrinsic Zeeman splitting in the GaAs. Note that circular dichroism is well avoided in this structure by the fact that the bandgap of the spin aligner is almost twice as large as the gap of the GaAs.

References

[1] G. A. Prinz, Phys. Today **48**, 58 (1995).

[2] F. G. Monzon, Mark Johnson, and M. L. Roukes, Appl. Phys. Lett. **71**, 3087 (1997).

[3] W. Y. Lee, S. Gardelis, B. C. Choi, Y. B. Xu, C. G. Smith, C. H. W. Barnes, D. A. Ritchie, E. H. Linfield, and J. A. C. Bland, J. Appl. Phys. **85**, 6682 (1999).

[4] S. Datta and B. Das, Appl. Phys. Lett. **56**, 665 (1990).

[5] G. Schmidt, L. W. Molenkamp, A. T. Filip, and B. J. van Wees, cond-mat 9911014 (1999).

[6] A. G. Aronov and G. E. Pikus, Sov. Phys. Semicond. **10**, 698 (1976); **15**, 1215 (1997).

[7] B. Beschoten, P. A. Crowell, I. Malajovich, D. D. Awschalom, F. Matsukura, A. Shen, and H. Ohno, Phys. Rev. Lett. **83**, 3073 (1999).

Metal–insulator transition in Si–MOS structures

A. Prinz, V. M. Pudalov, G. Brunthaler, G. Bauer[†]

Institut fur Halbeiterphysik, Johannes Kepler Universitat, Linz, A 4040, Austria

A survey is presented on recent investigations of the metal-to-insulator transition in two-dimensional systems with special emphasis on n-Si–MOS structures. Experimental facts are presented and the currently open questions on the nature of this transition are addressed.

© 2000 Academic Press

Key words: electronic phase transition, weak localization, magnetoconductance.

1. Introduction

The experimental evidence for a metal–insulator transition in two-dimensional (2D) electron systems, found by Kravchenko *et al.* [1] a few years ago, has created a widespread interest in this topic. The early observations on n-channel Si–MOS transistors were followed by experiments on 2D electron and hole systems in Si/SiGe heterostructures, and also on holes and electrons in GaAs/GaAlAs structures. In all systems evidence for a metallic behavior was found. The early findings were subjected to a lot of scepticism, mainly due to the fact, that in 1979 a seminal paper on the one-parameter scaling theory on transport was published by Abrahams *et al.* [2], according to which the conductance in a 2D system vanishes for temperatures approaching zero. The experimental observations of a more or less pronounced decrease of the resistivity with decreasing temperature signaling the presence of the metal-to-insulator transition are found for sample parameters where the Coulomb interaction between the carriers is sufficiently strong, i.e. where the Coulomb energy is substantially larger than the mean kinetic energy of a carrier ($r_s > 1$, where r_s is the ratio of Coulomb interaction energy E_{ee} to Fermi energy E_F), which lead to the assumption that carrier–carrier interactions are of decisive importance for the observation of the metal-to-insulator transition in 2D systems. Another equally important finding was the observation that magnetic fields applied parallel to the plane in which the 2D carriers are confined, suppress the anomalous metallic phase. This lead to the belief on the important role of the carrier spin for this type of MIT. Early reviews on general aspects of the MIT can be found in [3–5].

In the following we outline briefly the main physical considerations with respect to the MIT in 2D systems. In the one-parameter scaling theory of localization (OPST) by Abrahams *et al.* [2] the eigenfunctions of a quantum particle in a random potential are always localized when the number of dimensions is two or less. Therefore, the conductivity of infinite 1D and 2D conductors with arbitrarily weak disorder should vanish with temperature, at least in the absence of interactions between electrons. For finite 2D systems, the theory also describes a crossover from weak localization (WL), when the localization length ξ is much larger than the system size L, to strong localization (SL), when $\xi \ll L$. In the scaling theory, the temperature dependence

---

[†]E-mail: g.bavrer@hlphy.vni-linz.ac.at

of the conduction is governed by quantum corrections which develop as the length scale increases, which is determined by phase breaking length $l_\varphi = \sqrt{D\tau_\phi}$.

According to the OPST, there is no true metallic conduction and thus no metal–insulator transition possible for dimensions lower than 3. Until recently, the WL theory supplemented by a perturbative treatment of the electron–electron interaction had enjoyed a convincing agreement with numerous experiments. This agreement strengthened the belief that the scaling theory's description of WL and SL as well as the WL to SL crossover is qualitatively correct, even in the presence of electron–electron interactions.

The calculation of interaction quantum corrections meets essential difficulties for $r_s > 1$. However, with progress in the growth and fabrication of cleaner and cleaner 2D systems, the measurements entered into the regime of r_s as large as 10. The efforts to overcome substantial problems in the theoretical consideration of the scaling theory of the conductance were done by Finkelstein [6], Castellani *et al.* [7], Varma *et al.* [8], Belitz *et al.* [9], and resulted in the two-parameter scaling theory (TPST). In the TPST, not only the conduction, but primarily interaction itself is renormalized as the length scale grows (or temperature decreases). The two-parameter renormalization group equations predict a strong growth of the effective interaction constants, the process which finally should lead to an instability and magnetic phase transition in the electron spin subsystem. Concomitantly, the resistivity was predicted first to grow as T decreases (from the Drude limit) and then to drop; for lower temperatures the TPST assumptions become invalid. The applicability of this picture to the experimental systems under studies remains unclear so far.

The first indications that the prediction of the WL theory of insulating states in 2D may not be universally valid were obtained in studies of the quantum Hall effect [1, 11]. Extended states, which in high magnetic field B are located at centers of Landau bands, were found experimentally to merge and remain in a finite energy range as $B \to 0$. This behavior is not expected within the scaling theory, which predicts that the extended states 'float up' in energy in that limit [12, 13]. A serious challenge to the scenario of the insulating ground state in 2D arose in 1994 and the following years, when metallic behavior of electrical transport in zero magnetic field was observed in low-density and ultra-high mobility Si MOSFET structures [14–16]. Subsequently, qualitatively but not quantitatively similar metallic states were found in other 2D systems, including Si/Ge [17, 18], p-GaAs [19–22], n-GaAs [23, 24], and n-AlAs [25].

The two important features of the experimental findings are as follows:

(1) If the density of 2D carriers n is larger than some critical value n_c, the temperature T dependence of the resistance ρ is *metallic-like* [26]

$$\rho(T) = \rho_0 + \rho_1 \exp(-T_0(n)/T)^p \tag{1}$$

with $d\rho/dT > 0$, $p \sim 1$. The change in ρ can be as large as an order of magnitude in Si MOSFET [14, 26, 27] or as small as 10% in n-GaAs, [23].

(2) When $n < n_c$, both the sign ($d\rho/dT < 0$) and the functional (exponential) form of the $\rho(T) = \exp(T_0/T)^p$ dependence is characteristic of the insulating behavior.

2. Strong temperature dependence of the resistivity. Phase transition, quantum or quasiclassical scattering effect?

The exponential resistance drop as temperature decreases is the most striking appearance of the 'metallic' phase. For its explanation, a number of models were put forward, which are based on (i) Fermi liquid and (ii) nonFermi-liquid approaches. In the Fermi-liquid approach, the strong exponential increase of the conductivity, as the temperature decreases, is considered to be a finite temperature effect and several suggestions were put forward for its explanation: (i) unconventional spin-dependent carrier scattering [26], (ii) carrier localization and interaction with traps at the interface [28–31], or to (iii) disorder induced renormalization of the interaction [6–8, 10].

Table 1: Comparison of different material systems with respect to parameters yielding an r_s value of 10.

Material	m^*	ε_r	n (10^{10} cm^2)
n-Si–MOS	0.19	7.8	6.7
p-SiGe/Si	0.3	12.6	6.4
n-Si/SiGe	0.19	11.7	3.0
p-GaAs/AlGaAs	0.32	10.9	9.8
n-GaAs/AlGaAs	0.067	10.9	0.43

We note that for a Fermi-liquid system, the ground state is anticipated to be a conventional weakly localized state, unless the system experiences a transition to a new phase. A number of nonFermi-liquid models were suggested including: (i) a carrier singlet and triplet superfluid state [32–35], and (ii) a perfect metal state [36] with infinite G in the $T = 0$ limit.

The mirror-reflection symmetry of the experimental data around the critical carrier density seems to be a strong hint for the metal–insulator transition (MIT) in two-dimensional systems. Figure 1 demonstrates temperature dependence of the resistivity in an n-Si–MOS sample for a wide density range from the insulating ($n < n_c$) to the 'metallic' ($n > n_c$) regime. The temperature is normalized by E_F in order to highlight the following important features.

(1) The strong exponential drop in resistivity is a finite temperature effect observed in the range of $k_B T = (0.1 - 0.5)E_F$ only. For higher temperatures, the system evidently enters the Drude regime, whereas for lower T's the mirror symmetry between insulating and 'metallic' regions is lost.

(2) The strong drop in resistivity is not only present in the range of densities around n_c, but is a rather generic property for very wide range of densities [37].

(3) At higher densities ($n > 20 \times n_c$), $d\rho(T)/dT$ is metallic (positive) for higher temperatures ($k_B T > 0.1E_F$) but ρ exhibits a small (logarithmic) increase at low temperatures $T < 0.01E_F$ as in WL [37–40]. Thus, the exponential drop in $R(T)$ and mirror-reflection symmetry cannot be considered as proof for the occurrence of a true metallic state in two dimensions. Another important conclusion may be drawn from this picture: the conductivity of the 2D system is *limited by a finite value* $G_{max} \sim 140e^2/h$ whenever either T decreases or n increases, leaving thus no room for occurrence of the perfect metal [37].

Since the exponential drop in resistivity shown in Fig. 1 extends to temperatures as high as 50 K, it is very unlikely to be associated with the onset of a superconductivity as proposed in [32, 33, 35].

As already mentioned, a key parameter for the 2D systems exhibiting the MIT is the strength of the Coulomb interaction characterized by r_s. In order to compare different materials, in Table 1 the parameters are given for which r_s values of 10 are obtained. It can be seen that in n-GaAs/GaAlAs a very small carrier density of 4.3×10^9 cm^{-2} is required.

3. Phonon scattering, intersubband scattering; quantum versus quasiclassical origin of the strong drop in resistance

In the GaAs/AlGaAs system, due to the strong piezo-coupling, the contribution of electron–phonon interaction to the carrier scattering is not negligibly small. It was demonstrated in [22] that phonon scattering accounts for a maximum in $R(T)$ at temperatures $k_B T > 0.5E_F$. For n-type systems, especially, for Si lattices with rather weak deformation potential coupling, the phonon contribution is much less; calculations [41]
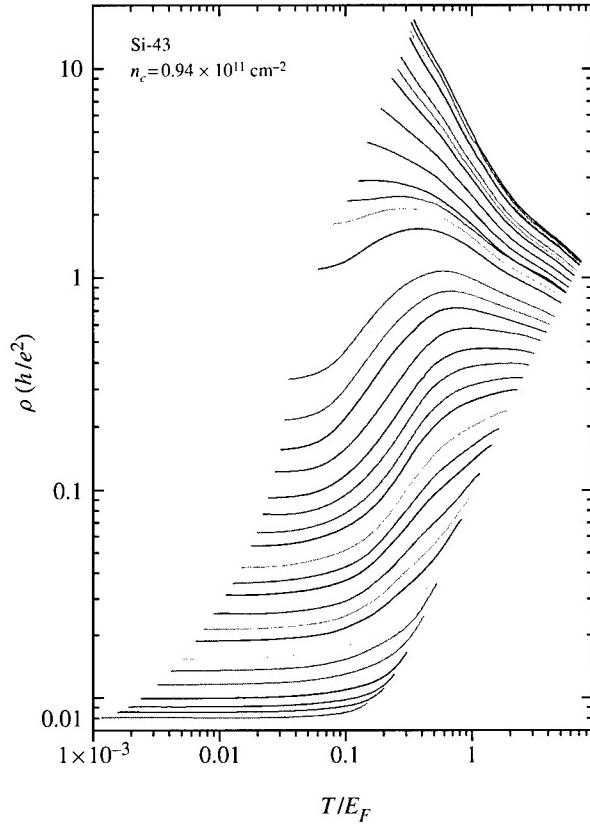
Fig. 1. Resistivity versus temperature (from 0.3 to 45 K) for a wide range of density and Fermi energy. Density values are between 0.707 (top) and 35×10^{11} cm^{-2} (bottom).

show that its contribution does not exceed 1% for temperatures of ≈ 50 K and raises to 10% in the limit of very high densities ($n \geq 50 \times 10^{11}$ cm^{-2}) and higher temperatures only.

In the p-type GaAs/AlGaAs system the lowest heavy hole subband is degenerate and splits for higher values of the in-plane k-vector in two bands. Consequently, inelastic scattering between these states is possible for sufficiently high hole concentrations. Yaish *et al.* [42] have attributed the resistance decrease with decreasing temperature to inelastic interband scattering of the heavy holes, caused by acoustic plasmon mediated Coulomb scattering. This results in the temperature-dependent drop of the resistivity by a factor of 2 [42]. Such an interpretation can also be verified by measurements of the magnetoresistance in a weak perpendicular field at a constant low temperature. Indeed in p-type GaAs/GaAlAs a *positive* magnetoresistance with a Lorentzian shape and of the same magnitude as the above-mentioned resistance drop with temperature is observed. According to this interpretation, the effect of the strong temperature-dependent resistance drop in p-GaAs/AlGaAs (by a factor of two) is therefore of classical origin and caused by interband scattering of carriers between the split heavy hole states of the lowest subband at finite k-values. For even lower temperatures, the weak localization logarithmic quantum correction was found in p-GaAs/AlGaAs [39, 40] to set in, replacing the 'metallic'-like T-dependence.

In n-Si–MOSFETs, the two-electron valleys are located rather far away in the Brillouin zone, close to X-points, and the intervalley scattering rate is negligible at low temperatures. In addition, for carrier concentrations up to about $n = 5 \times 10^{12}$ cm^{-2} the second electron subband is not occupied at low temperatures.

In agreement with this properties of the Si–MOS system, no positive magnetoresistance is observed in a weak perpendicular field. Instead, a regular negative magnetoresistance occurs, which can be fitted with the digamma functional shape [15,43]; the interpretation of the weak field magnetoresistance as being due to an interference correction is supported by the corresponding small amplitude of the magnetoresistance $\delta\rho/\rho \sim 1$–3% and its sensitivity to only the perpendicular component of the magnetic field.

The effect of the strong drop in resistance takes place as mentioned above, in the range of temperatures, $k_B T \sim (0.1$–$0.5 E_F)$. We have investigated the negative magnetoresistance due to the weak-localization effect for high electron densities in order to get information about the phase coherence time τ_φ [43]. Figure 2 shows magnetoresistance measurements on sample Si-43 for an electron density $n = 2.6 \times 10^{12}$ cm^{-2} for temperatures between 0.29 and 10.7 K. The measured data are represented by the points whereas the full line shows the best least square fit of the magnetic field dependence according to the single-electron weak localization correction to conductivity [45]

$$\Delta\sigma_{xx} = -\frac{\alpha g_\nu e^2}{2\pi^2 \hbar} \left[\Psi \left(\frac{1}{2} + \frac{\hbar}{4eBD\tau} \right) - \Psi \left(\frac{1}{2} + \frac{\hbar}{4eBD\tau_\varphi} \right) \right], \tag{2}$$

where Ψ is the Digamma function, B is the applied perpendicular magnetic field, D is the diffusion coefficient and τ the momentum relaxation time. The values for D and τ were deduced from temperature and density dependent Hall and resistivity measurements. The prefactor g_ν describes the valley degeneracy and α depends on the ratio of intra-valley to inter-valley scattering rates and was found to be between 0.5 and 0.6. Figure 3 shows the temperature dependence of the phase coherence time τ_φ for several carrier densities in the range from 5.4×10^{11} to 3.5×10^{12} cm^{-2} for sample Si-43. It is found that τ_φ increases by about a factor of 100 from 1 ps at 10 K to nearly 100 ps at 0.29 K. This result shows that the quantum phase breaking time τ_φ becomes less than τ_p, the momentum relaxation time at temperatures of about 10 K or $\sim 0.1 E_F$. In other words, at high densities the strong drop of the resistivity develops entirely in the quasiclassical temperature and length scale range, where the quantum renormalization has not yet started. We conclude that both for the GaAs/GaAlAs and the Si–MOS systems, the remarkable effect of the strong drop in resistance is a quasiclassical or classical but not a quantum phenomenon.

4. Symmetry of the confining potential and the spin contribution

The main evidence for the importance of electron spins for the resistance drop is considered to be obtained from the strong increase of the resistivity in a parallel magnetic field [15,46,47]. A remarkable feature is highlighted in the behavior of $\rho(B_\parallel)$ as shown in Fig. 4 ([15]). The resistance in a Si–MOS sample, increases substantially with increasing magnetic field up to a factor of about 30, depending on the sheet carrier density. The resistance saturates at magnetic fields where the Zeeman energy $g\mu B$ equals the Fermi energy, as shown by Okamoto *et al.* [44]. This condition is equivalent to full polarization of the electron spins.

Several theoretical models for parallel field magnetoresistance have been proposed so far. Das Sarma *et al.* [57] recently reported that the strong magnetoresistance arises due to a transition from the 2D- to 3D-character of the electronic transport, when the magnetic length becomes less than the thickness of the 2D layer (the width of the confining potential in z-direction is about 100 Å). The predicted magnetoresistance is to be strongly anisotropic (by a factor of 2) with respect to the direction of the in-plane magnetic field along or perpendicular current. This strong anisotropy is indeed found for p-GaAs/AlGaAs [48], where the confining potential is rather soft, and is to a large extent determined self-consistently by the wavefunction of the charge carriers. However, in n-Si–MOS structures the anisotropy of the magnetoresistance was found [49] not to exceed 1–2% and thus cannot be associated with a 2D–3D transition.

Finally, the contribution of the symmetry breaking of the confining potential was predicted to account for a specific anisotropy of the magnetoresistance in a parallel field [50]. The anisotropy factor is small $(\rho_\parallel - \rho_\perp)/(\rho_\parallel + \rho_\perp) \sim 0.1\Delta_z/\alpha k_F \sim 1$–3% at most, where $\rho_\perp$ and $\rho_\parallel$ are the resistivity measured
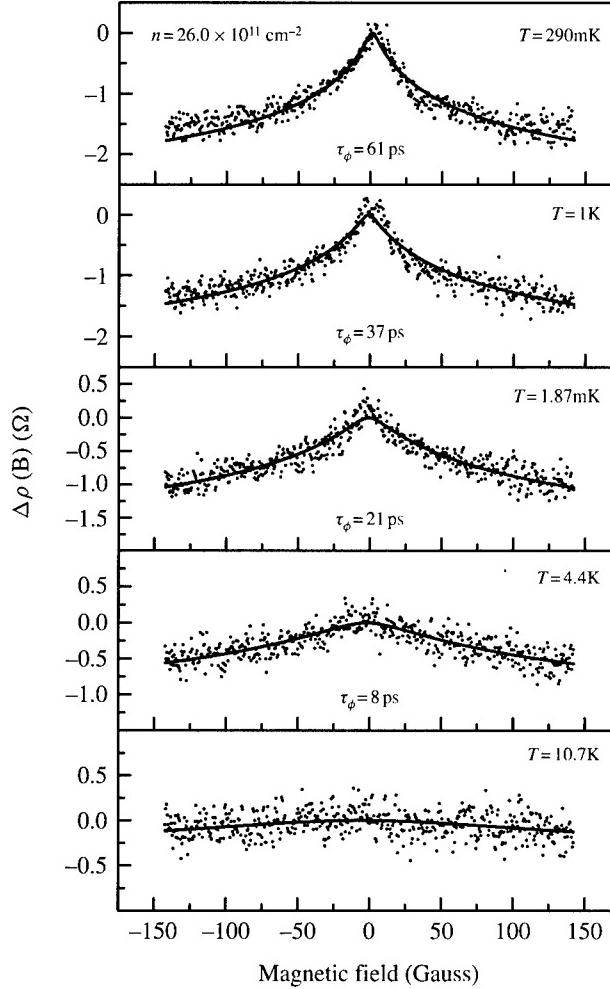
Fig. 2. Change of resistivity versus magnetic field for n-Si–MOS sample Si-43 at an electron density of 2.6×10^{12} cm^{-2} at several temperatures. The dots represent the measurements and the full lines the results of model calculations for the weak localization dependence.

across and along the in-plane magnetic field. This model predicts specifically, that the anisotropy factor is magnetic field dependent: it exhibits a sharp maximum at a field where the Zeeman energy Δ_Z is equal to spin splitting at the Fermi energy αk_F in the Rashba spectrum. Both of these predictions are seen in the measured magnetoresistance in a parallel field in Si–MOS structures [49]. Finally, in a weak perpendicular field in high mobility Si–MOS structures we found pronounced beatings of Shubnikov–de Haas oscillations, from which we calculate the corresponding difference in k_F vectors and the zero-field spin splitting αk_F. We conclude that the asymmetry of the confining potential in Si–MOS structures does indeed correlate with the strong drop in resistance and may account for it.
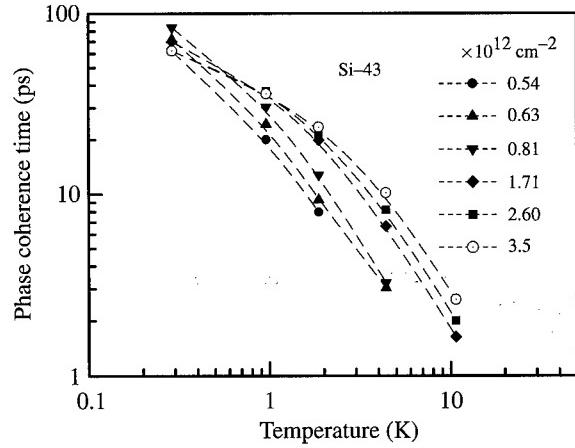
Fig. 3. Temperature dependence of the phase coherence time τ_ϕ for various 2D carrier densities for n-Si–MOS sample 43.
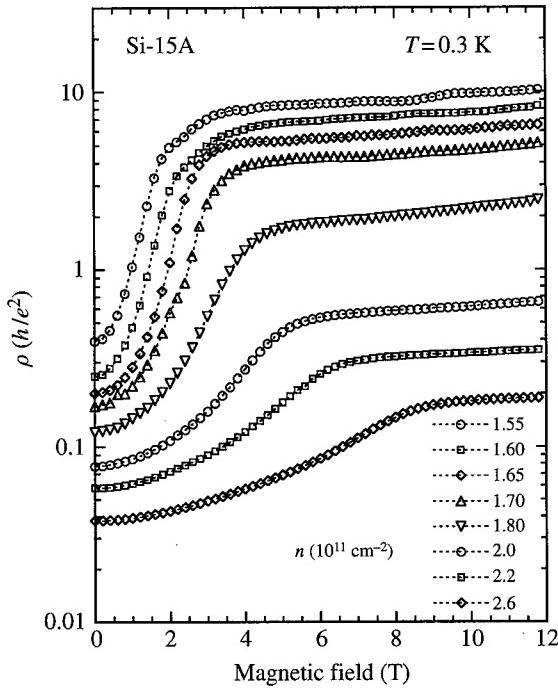
Fig. 4. Change of resistivity of an n-Si–MOS sample (Si-15A) as a function of the parallel magnetic field for electron various densities at $T = 0.3$ K.

5. Temperature dependence of the conductivity and the ground state of the system in the $T = 0$ limit

All data presented above are related to the most striking feature, the strong exponential-like increase of the conductivity (or drop in resistivity) as T decreases. It is clear from the above analysis that at least for two systems most intensively studied, n-Si–MOSFETs and p-GaAs/AlGaAs, the effect is not of a quantum

origin. However, the origin of this anomalous quasiclassical scattering remains so far unclear at least for the Si–MOS system.

In order to find experimentally the answer to the question on the nature of the ground state conduction, one should evidently make measurements in the range of sufficiently low temperatures, so that the influence of the quasiclassical scattering effects do not mask the quantum corrections. The latter ones were empirically found to occur for $k_B T < 0.007 E_F$ [27,37,38]. In the analysis of the data for high mobility n-Si–MOS structures, we used: (i) the logarithmic corrections from $R(T)$ data at zero field, (ii) logarithmic corrections $R(T)$ from the data measured at weak perpendicular field of 0.1 T, which suppresses the interference correction to the conductivity, and (iii) logarithmic temperature corrections to the Hall resistivity measured at field of 0.1 T. As a result, we found that the interaction correction has the same localizing sign as the single-particle interference correction. For the range of temperatures from $0.007 E_F$ to $0.1 E_F$ we found the presence of a linear temperature dependence $\rho(T) \propto k_B T / E_F$ predicted in [31,51–56]. This weak temperature dependence extends down to temperatures ($\sim 0.01 E_F$) much lower than the exponential one does, and its presence creates the illusion of the existence of a delocalization or a temperature-independent conduction when data are taken at not sufficiently low temperatures [61].

6. Conclusion: Mobility edge or Mott–Anderson transition at finite temperatures? Ground state conduction

In recent compressibility measurements [58,59], clear signatures of a step-like discontinuity in the chemical potential was found across the critical density n_c as obtained from the low-temperature transport data in p-GaAs/GaAlAs structures. These results seem to imply that the metallic and insulating phase is indeed revealed in a thermodynamic property. Consequently, besides the scattering between the split lowest heavy hole subband as found by Yaish *et al.* [42] in p-GaAs/GaAlAs another mechanism might contribute to the change of the conductance with temperature.

Similarly, in measurements of the thermoelectric power S performed across the critical density in n-Si–MOS structures [60], a sharp transition was found from metallic-like dependence $S \propto T$ (for $n > n_c$) to the divergence of S/T in the insulating regime. These results, in addition to the transport data, give a clear evidence for the mobility edge or Mott–Anderson transition being present at finite temperatures. However, for even lower temperatures this scenario fails; as soon as the length scale renormalization turns on, all the studied systems enter the WL regime, as anticipated for the regular Fermi liquid or the disordered metal.

It remains still a puzzle why the conductance behavior appears to be so general for several different material systems. Another intriguing question remains unanswered, whether or not the strong interaction scenario suggested by the two-parameter scaling theory may dominate the weak localization correction for sufficiently low temperatures $k_B T \ll 0.001 E_F$ in the metallic regime.

References

[1] S. V. Kravchenko, W. Mason, J. E. Furneaux, and V. M. Pudalov, Phys. Rev. Lett. **75**, 910 (1995).
[2] E. A. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).
[3] N. F. Mott and E. A. Davis, *Electron Processes in Non-crystalline Materials* (Clarendon Press, Oxford, 1979) 2nd edn.
[4] For a review see: P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).

[5] B. L. Altshuler and A. G. Aronov, *Electron-Electron Interactions in Disordered Systems*, edited by A. L. Efros and M. Pollak (North-Holland, Amsterdam, 1985).

[6] A. M. Finkel'stein, Z. Phys. **B56**, 189 (1984); Sov. Sci. Rev., Ed. Khalatnikov, I. M. **14**, 3 (1990).

[7] C. Castellani, C. DiCastro, H. P. Fukuyama, A. Lee, and M. Ma, Phys. Rev. **B33**, 7277 (1986); C. Castellani, C. DiCastro, and P. A. Lee, cond-mat/9801006.

[8] Q. Si and C. M. Varma, Phys. Rev. Lett. **81**, 4951 (1998).

[9] D. Belitz and T. R. Kirkpatrick, Rev. Mod. Phys. **66**, 261 (1994).

[10] X. Waintal, G. Benenti, and J. L. Pichard, cond-mat/9906397.

[11] M. D'Iorio, V. M. Pudalov, and S. G. Semenchinsky, Phys. Lett. **A150**, 422 (1990); Phys. Rev. **B46**, 15992 (1992); V. M. Pudalov, M. D'Iorio, and J. W. Campbell, JETP Lett. **57**, 608 (1993); Surf. Sci. **305**, 107 (1994).

[12] D. E. Khmelnitskii, Phys. Lett. **106A**, 673 (1984).

[13] R. B. Laughlin, Phys. Rev. Lett. **52**, 2304 (1984).

[14] S. V. Kravchenko, G. V. Kravchenko, J. E. Furneaux, V. M. Pudalov, and M. D'Iorio, Phys. Rev. **B50**, 8039 (1994); S. V. Kravchenko *et al.*, Phys. Rev. **B51**, 7038 (1995).

[15] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, JETP Lett. **65**, 932 (1997).

[16] D. Popovi'c, A. B. Fowler, and S. Washburn, Phys. Rev. Lett. **79**, 1543 (1997).

[17] P. T. Coleridge, R. L. Williams, Y. Feng, and P. Zawadzkii, Phys. Rev. **B56**, R12764 (1997).

[18] J. Lam, M. D'Iorio, D. Brown, and H. Lafontaine, Phys. Rev. **B56**, R12741 (1997); D'Iorio *et al.*, Superlatt. Microstruct. **23**, 55 (1998).

[19] M. Y. Simmons, A. R. Hamilton, M. Pepper, E. H. Linfeld, P. D. Rose, D. A. Ritchie, A. K. Savchenko, and T. G. Griffiths, Phys. Rev. Lett. **80**, 1292 (1998).

[20] Y. Hanein, U. Meirav, D. Shahar, C. C. Li, D. C. Tsui, and H. Shtrikman, Phys. Rev. Lett. **80**, 1288 (1998).

[21] S. J. Papadakis *et al.*, Science **283**, 2056 (1999).

[22] A. P. Mills, A. P. Ramirez, L. N. Pfeiffer, and K. W. West, cond-mat 9905176.

[23] Y. Hanein, D. Shahar, C. C. Li, D. C. Tsui, and H. Strikman, Phys. Rev. **B58**, R13338 (1998).

[24] E. Ribeiro, R. D. Jäggi, T. Heinzel, K. Ensslin, G. Medeiros-Ribeiro, and P. M. Petroff, Phys. Rev. Lett. **82**, 996 (1999).

[25] S. J. Papadakis and M. Shayegan, Phys. Rev. **B57**, R15068 (1998).

[26] V. M. Pudalov, JETP Lett. **66**, 170 (1997); [Pis'ma ZhETF **66**, 168 (1997)].

[27] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, JETP Lett. **68**, 442 (1998); [Pis'ma ZhETF **68**, 415 (1998)].

[28] B. L. Altshuler and D. L. Maslov, Phys. Rev. Lett. **82**, 145 (1999).

[29] B. L. Alsthuler, D. L. Maslov, and V. M. Pudalov, cond-mat/9909353.

[30] T. M. Klapwijk and S. Das Sarma, Solid State Commun. **110**, 1581 (1999).

[31] S. Das Sarma and E. H. Hwang, Phys. Rev. Lett. **83**, 164 (1999).

[32] P. Phillips *et al.*, Nature **395**, 253 (1998).

[33] F.-C. Zhang and T. M. Rice, cond-mat/9708050.

[34] C. Castellani, C. Di Castro, and P. A. Lee, Phys. Rev. **B57**, R9381 (1998).

[35] D. Belitz and T. R. Kirkpatrick, Phys. Rev. **B58**, 8214 (1999).

[36] V. Dobrosavljević, E. Abrahams, E. Miranda, and S. Chakravarty, Phys. Rev. Lett. **79**, 455 (1997).

[37] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, Phys. Rev. **B60**, R2154 (1999).

[38] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, JETP Lett. **68**, 534 (1998); [Pis'ma ZhETF **68**, 497 (1998)].

[39] A. R. Hamilton, M. Y. Simmons, M. Pepper, E. H. Linfield, P. D. Rose, and D. A. Ritchie, cond-mat/9808108.

[40] M. Y. Simmons *et al.*, cond-mat/9910368.

[41] D. L. Maslov, B. L. Altshuler, and V. M. Pudalov, cond-mat/0003032.

[42] Y. Yaish, O. Prus, E. Buchstab, S. Shapira, G. Ben Yosef, U. Sivan, and A. Stern, cond-mat/9904324.

[43] G. Brunthaler, A. Prinz, G. Bauer, V. M. Pudalov, E. M. Dizhur, J. Jaroszynski, P. Glod, and T. Dietl, Ann. Phys. **8**, 579 (1999).

[44] T. Okamoto, K. Hosoya, S. Kawaji, and A. Yagi, Phys. Rev. Lett. **82**, 3875 (1999).

[45] B. L. Altshuler, D. E. Khmelnitskii, A. I. Larkin, and P. A. Lee, Phys. Rev. **B22**, 5142 (1980); H. Fukuyama, Surf. Sci. **113**, 489 (1982).

[46] D. Simonian, S. V. Kravchenko, M. P. Sarachik, and V. M. Pudalov, Phys. Rev. Lett. **79**, 2304 (1997).

[47] J. Yoon, C. C. Li, D. Shahar, D. C. Tsui, and M. Shayegan, cond-mat/9907128.

[48] S. J. Papadakis, E. P. De Poortere, and M. Shayegan, cond-mat/9911239.

[49] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, cond-mat/0004206.

[50] G. H. Chen, M. E. Raikh, and Y. S. Wu, cond-mat/9904451.

[51] F. Stern, Phys. Rev. Lett. **44**, 1469 (1980).

[52] A. Gold and V. T. Dolgopolov, Phys. Rev. **B33**, 1076 (1986).

[53] K. K. Choi, Phys. Rev. **B28**, 5774 (1983).

[54] S. Das Sarma, Phys. Rev. **B33**, 5401 (1986).

[55] E. A. Vyrodov, V. T. Dolgopolov, S. I. Dorozhkin, and N. B. Zhitenev, ZhETF **94**, 234 (1988); [Sov. Phys. JETP **67**, 998 (1998)].

[56] K. M. Cham and R. G. Wheeler, Phys. Rev. Lett. **44**, 1472 (1980).

[57] S. Das Sarma and E. H. Hwang, cond-mat/9909452.

[58] S. Ilani, A. Yacoby, D. Mahalu, and H. Shtrikman, cond-mat/9910116.

[59] S. C. Dultz and H. W. Jiang, cond-mat/9909314.

[60] T. Fletcher, A. D. B. Radcliffe, V. M. Pudalov, and C. Possanizini, cond-mat/0002436.

[61] S. V. Kravchenko and T. Klapwijk, cond-mat/9909458.

Metal–insulator transition in quantum dot arrays

A. Shailos, M. El Hassan, C. Prasad, J. P. Bird, D. K. Ferry

Department of Electrical Engineering and Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85287-5706, U.S.A.

L.-H. Lin, N. Aoki, K. Nakao, Y. Ochiai

Department of Materials Science, Chiba University, 1-33 Yayoi, Inage, Chiba 236-8522, Japan

K. Ishibashi, Y. Aoyagi, T. Sugano

Nanoelectronic Materials Laboratory, Frontier Research Program, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

We present evidence for a re-entrant metal–insulator transition that arises in quantum dot arrays as the gate voltage is used to sweep their density of states past the Fermi level. The form of the temperature variation of the conductance observed in these arrays can be accounted for using a functional form derived from studies of the metal–insulator transition in two dimensions, although the values obtained for the fit parameters suggest that the behavior we observe here may be quite distinct to that found in two dimensions.

© 2000 Academic Press

Key words: localization, metal–insulator transition, quantum dots.

The existence of a metal–insulator transition in two dimensions has now been confirmed in a number of experiments, performed on a variety of semiconductor systems [1]. In this report, however, we provide experimental evidence for a metal–insulator transition in a system of even lower dimensionality, namely a linear array of coherently coupled quantum dots. These arrays are realized using the split-gate technique and no evidence for the metal–insulator transition is found prior to biasing the gates to form the arrays. With a suitable voltage applied to the gates to form the array, its resistance is typically found to diverge with decreasing temperature, which behavior is characteristic of localization. At a small number of isolated gate voltages, however, the resistance actually decreases logarithmically with temperature, indicating that a transition to a metallic state has occurred. Based on studies of the metal–insulator transition in two dimensions, we are able to propose a simple functional form that accounts reasonably well for the temperature-dependent variation of conductance observed in many cases. We also propose a basic mechanism for the implied metal–insulator transition, which appeals to the discrete form of the density of states in the mesoscopic arrays.

Linear arrays consisting of three identical quantum dots connected in series were realized using the split gate technique [2]. Au–Ti gates were deposited on GaAs/AlGaAs heterojunction Hall bars, whose low-temperature carrier density was $3.8 \times 10^{11} \mathrm{cm}^{-2}$ and whose mobilities ranged from 2×10^5 to $10^6 \mathrm{cm}^2 \mathrm{V}^{-1} \mathrm{s}^{-1}$.
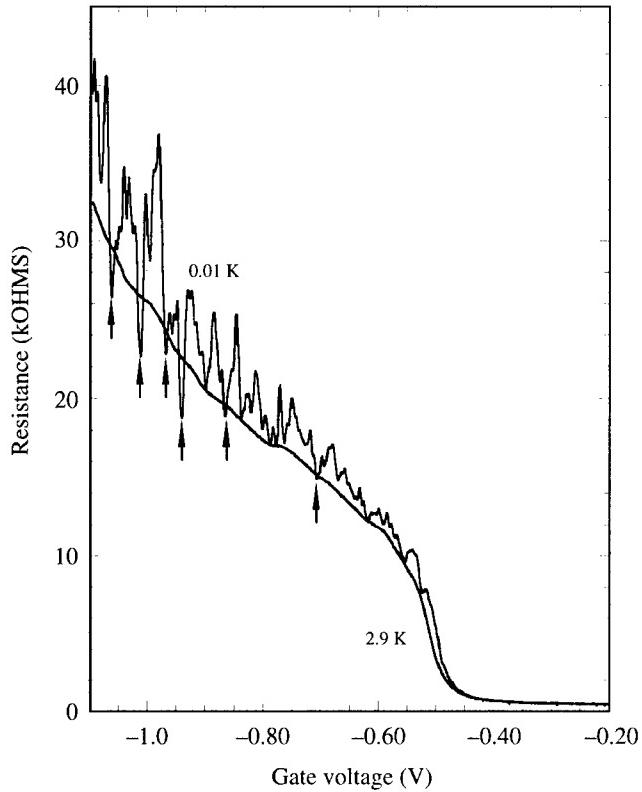
Fig. 1. Resistance-gate voltage characteristic of a quantum dot array composed of three dots with lithographic dimensions of 0.4 × 0.7 μm.

Five different arrays were investigated for this study, and were all found to reveal qualitatively similar behavior. In three of the arrays, the lithographic size of each of the dots was 1.0 × 0.6 μm, while the other two arrays were constructed from dots with dimensions of 1.3 × 0.8 μm and 0.7 × 0.4 μm. The key details of the behavior we discuss here were found to be reproduced after thermally cycling the devices to room temperature and after low-temperature illumination with a red LED. An analysis of the magneto-conductance of the arrays revealed the highly coherent nature of electron transport through them, with a phase breaking time of order 100 ps at the lowest temperatures and a corresponding ballistic path length of more than 30 μm [3]. The arrays were mounted in good thermal contact with the mixing chamber of a Kelvinox 100 dilution refrigerator and measurements of their linear conductance were made using small enough currents to avoid electron heating [3–5].

In Fig. 1 we show the results of measurements of the resistance-gate voltage characteristic of one of the arrays. The reproducible fluctuations, found to emerge on cooling to the milli-Kelvin range, are thought to result from a density of states variation that is induced as the dot size is modulated by the gate voltage [6]. Of interest here, however, is the existence of a small number of gate voltage values at which a metal–insulator transition appears to occur. As can be seen from Fig. 1, the general trend on decreasing temperature is for an increase in resistance, behavior that is characteristic of localization. At a limited number of gate voltages (marked by the arrows), however, the low-temperature value of the resistance is seen to drop below the higher-temperature curve, a variation that is indicative of a *metallic* state. Similar behavior to that shown in Fig. 1 was found in measurements performed on a number of different arrays. These measurements reveal

A   B
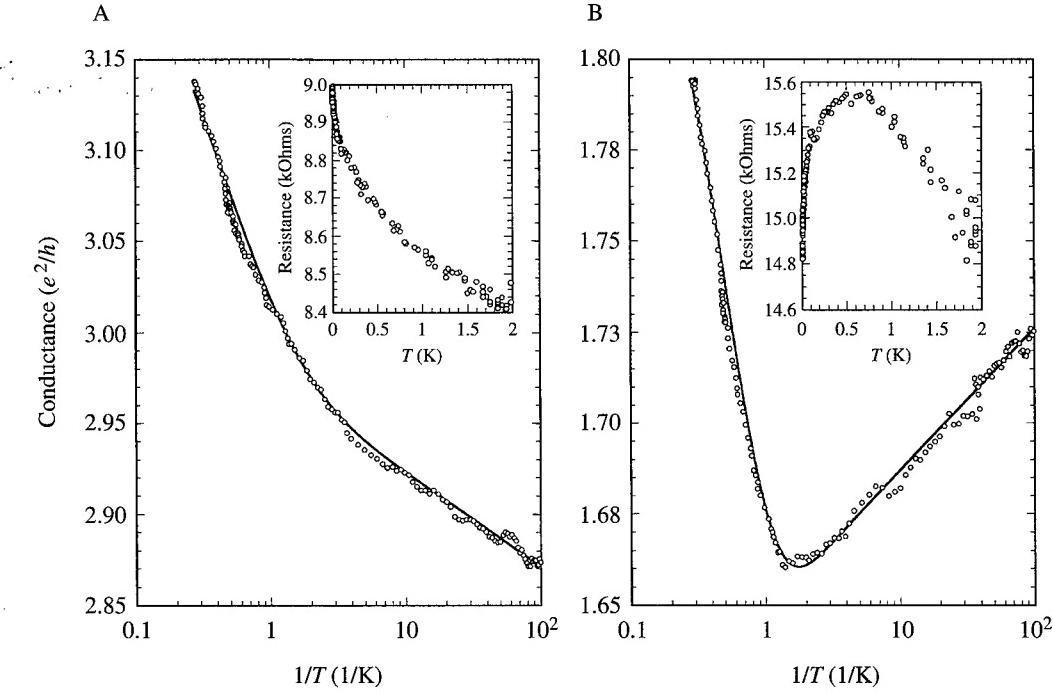
Fig. 2. Conductance variation with temperature, measured in a quantum dot array composed of three dots with lithographic dimensions of 1.3×0.8 μm. A, The gate voltage applied to the array in this case was adjusted such that the resistance corresponded to a local maximum. The solid line is a fit to the form of eqn (1a) with the parameters $G_0 = 2.975$ e^2/h, $G_1 = 0.0225$ e^2/h, $G_2 = 0.325$ e^2/h, $T_0 = 3.2$ K, and $p = 0.6$. The inset shows the corresponding variation of *resistance* with temperature. B, The gate voltage applied to the array in this case was adjusted such that the resistance corresponded to a local maximum. The solid line is a fit to the form of eqn (1b) with the parameters $G_0 = 1.648$ e^2/h, $G_1 = 0.017$ e^2/h, $G_2 = 0.28$ e^2/h, $T_0 = 2$ K, and $p = 1.2$. The inset shows the corresponding variation of *resistance* with temperature.

no obvious evidence for a dependence of the number of the metallic 'dips' on the component dot size, but instead suggest that this number is determined by the degree of disorder in the device. Indeed, studies of highly disordered arrays show no evidence for the metallic state whatsoever [2].

In order to more clearly demonstrate the observation of a metal–insulator transition, in Fig. 2 we show the results of temperature-dependent measurements of the conductance, at gate voltages corresponding to a local resistance maximum (Fig. 2A) and a local resistance minimum (Fig. 2B). Clearly, in the case of a local resistance maximum, the conductance decreases with increasing temperature, while for the local minimum the opposite behavior is found. The solid lines in Fig. 2 are functional fits to the forms [2,5]:

$$G_{ins}(T) = G_0 + G_1 \ln T + G_2 \exp\left[-\frac{T_0}{T}\right]^p \tag{1a}$$

$$G_{met}(T) = G_0 - G_1 \ln T + G_2 \exp\left[-\frac{T_0}{T}\right]^p. \tag{1b}$$

In these equations, G_{ins} is the conductance measured in the insulating regime while G_{met} is that measured in the metallic regime. The motivation for these forms is discussed in detail in [2]. Basically, however, the exponential term is thought to represent the excitation of carriers across some characteristic energy gap ($k_B T_0$). The main difference between the conductance in the insulating and metallic regimes is given by the *sign* of the logarithmic term. This term dominates the low-temperature (mK) behavior while at intermediate temper-

atures ($T \approx T_0$) an exponential decrease of conductance with decreasing temperature is found (see Fig. 2A and B). Quite generally, we find the value of the parameter p to be of order 1.5, which is considerably larger than the value of $1/2$ expected for variable-range hopping in the presence of a Coulomb gap [7]. One possibility is that this disagreement reflects the reduced dimensionality of the quantum dot arrays we study, although theoretical studied are required to confirm this possibility. Alternatively, the poor agreement of the power law exponent with the expected value of $1/2$ may indicate that variable-range hopping is not responsible for the temperature-dependent variation of the resistance in the 'exponential' regime. As for the metallic-like state, this appears to be characteristic of low-temperatures and one possibility is that its onset may be related to the corresponding growth of the electron–electron interaction. Another feature of Fig. 1 to note is that, simply by variation of the gate voltage, it appears to be possible to cycle successively between metallic and insulating states. Since the gate voltage is known to sweep the discrete states of the dot past the Fermi surface [6], the suggestion is that the localized or metallic behavior arises as localized or extended regions of the density of states are swept past the Fermi surface.

In conclusion, we have presented evidence for a re-entrant metal–insulator transition that arises in quantum dot arrays as the gate voltage is used to sweep their density of states past the Fermi level. The form of the temperature variation of the conductance observed in these arrays can be accounted for using a functional form derived from studies of the metal–insulator transition in two dimensions, although the values obtained for the fit parameters suggest that the behavior we observe here may be quite distinct to that found in two dimensions. Further theoretical and experimental studies are required to clarify the nature of the novel metallic state that appears to form in this mesoscopic systems.

References

[1] For recent developments see the *Proceedings of the International Conference on Localization 1999*, Ann. Phys. **8**, (1999) pp. SI-1–SI-300.
[2] A. Andresen *et al.*, Phys. Rev. **B60**, 16050 (1999).
[3] C. Prasad *et al.*, SIMD'99, Superlatt. Microstruct. **27**, 315 (2000).
[4] L.-H. Lin *et al.*, SIMD'99, Superlatt. Microstruct. **27**, 327 (2000).
[5] L.-H. Lin *et al.*, Phys. Rev. **B60**, R16299 (1999).
[6] J. P. Bird *et al.*, Phys. Rev. Lett. **82**, 4691 (1999).
[7] A. L. Efros and B. I. Shklovskii, J. Phys. **C8**, L49 (1975).

Phase-breaking time variations with temperature and current in an open quantum dot array

C. Prasad, A. Andresen, F. Ge, J. P. Bird, D. K. Ferry

Department of Electrical Engineering and Center for Solid State Electronics Research, Arizona State University, Tempe AZ 85287-5706, U.S.A.

L. H. Lin, N. Aoki, K. Nakao, Y. Ochiai

Department of Materials Science, Chiba University, 1-33 Yayoi, Inage, Chiba 236-8522, Japan

K. Ishibashi, Y. Aoyagi, T. Sugano

Nanoelectronic Materials Laboratory, Frontier Research Program, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

Magnetic characterization of a quantum dot array was carried out to extract the phase-breaking time as a function of temperature and current. These measurements indicate a saturation of the phase-breaking time at low temperature and low current. The phase-breaking time varies as $1/T$ for temperatures above 1 K, which can be attributed to carrier–carrier scattering processes. Variation of the phase-breaking time with current yields an estimate for the electron temperature.

© 2000 Academic Press

Key words: magneto-transport, heterostructures, phase-breaking, coherent transport.

As the electron temperature increases, the basic resistance of a mesoscopic structure changes as the mobility of the electrons is reduced. This causes decay in the phase-coherence time or the phase-breaking time of the device. It has been observed in previous experimental work that the phase-breaking time, τ_ϕ, exhibits a power law variation [1, 2] with temperature at temperatures above a few hundred mK and a saturation is observed in the value at lower temperatures, which is not associated with a corresponding saturation in the sample temperature and is, therefore, intrinsic to the system. This saturation is not predicted in present theories that do not take carrier–carrier scattering into account. However, this behavior is expected in both 2D and 1D due to electron–electron scattering.

The device studied is a quantum dot array consisting of four dots in series, with lithographic dot sizes of $1.0 \times 0.6 \ \mu\text{m}^2$ as shown in Fig. 1. The material used is a modulation-doped AlGaAs/GaAs heterostructure with a mobility of 243000 cm^2 V^{-1} s^{-1}, a zero field resistance of 483 Ω a carrier density of 3.68×10^{11} cm^{-2}. This gives a momentum relaxation time of 9.1 ps. Split-gate pairs were connected together to achieve even depletion. Characterization of this structure was carried out in a dilution refrigerator system. The magnetic field was swept over a range of 0 to 4 T and the corresponding magneto-conductance oscillations were recorded over a temperature range of 0.01–4.2 K with a constant sample current of 0.5 nA and over a current range of 0.5–280 nA with a constant sample temperature of 0.01 K. The recorded traces were processed by performing background subtraction and computation of a correlation function [3]. The critical magnetic field
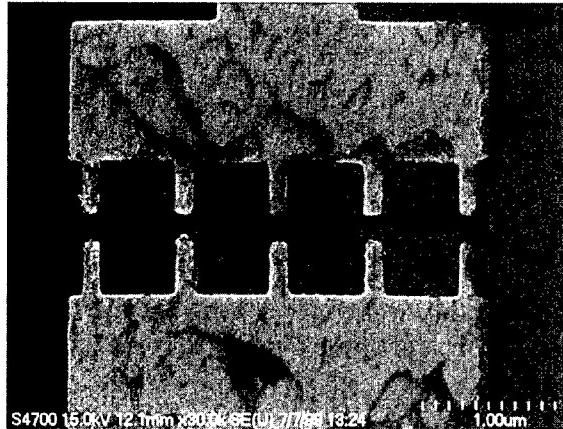
Fig. 1. SEM micrograph of the quantum dot array with four dots. The width of the quantum point contacts is 0.15 μm and the lead opening is 0.2 μm. Since the phase-breaking length computed for such a device on the basis of the measured values of the phase-breaking time is larger than its length, we must consider this device as a quantum dot array and not a quantum wire with constrictions.

as a function of the magnetic field was determined from the correlation functions and the phase-breaking time was extracted by linear fits [4] to the critical magnetic field. The latter quantities are given by:

$$B_c(B) = \frac{8\pi^2 m^* B}{\hbar k_F^2 \tau_\phi}.$$

Estimation of the electron temperature was carried out by equating the resistance-temperature and the resistance-current characteristics and extracting a current-temperature characteristic, to which the phase-breaking time data was fitted. Also, the energy relaxation time was determined by using the current-temperature characteristic to extract the power dissipated per electron at a given temperature.

In 3D and 2D systems, the Hartree approximation shows that the electron–electron relaxation time, τ_{ee}, varies as T^{-1} under degenerate conditions and is independent of temperature since the interaction energy range is the energy itself in 1D and 0D systems. From our experimental data shown in Fig. 2, we see that the decay of the phase-breaking time can be fit very well to the predicted T^{-1} behavior at higher temperatures. Such behavior has been previously observed by Bird *et al.* [5] and been attributed to a possible dimensional crossover. However, in the present experiments, saturation of the phase-breaking time occurs in the temperature range of 750–1050 mK whereas the predicted crossover temperature according to the dot area and energy level spacing is 132 mK, which is a factor of six smaller than the observed value. We also note that as the gate voltage is decreased, thereby increasing the dot size, the phase-breaking time decreases, corresponding to an inverse size scaling effect. Pivin *et al.* [6] reported a possible explanation for this behavior, where it is suggested that the phase-breaking interaction is a function of the total number of electrons in the dot. Measurements carried out by Marcus *et al.* [7] and Clarke *et al.* [8] using the power spectrum of the magneto-conductance fluctuations to determine the phase-breaking time yield similar results. The values of the phase-breaking time measured in the present experiment are a factor of three lower than other experimental measurements and this could be due to some superlattice effect of the dot array.

At low current values, we see that the phase-breaking time exhibits a saturation similar to that exhibited by it with variations in temperature. However, at higher values of current, carrier heating takes place and this leads to a rise in the electron temperature of the system and thereby causes the phase-breaking time to decay. If electron–electron scattering is the only dominant process, then the rate of decay of the phase-breaking time with lattice temperature should be the same as that with electron temperature [9]. The plots of phase-breaking time with electron temperature show that the decay rate agrees very well with that of electron–
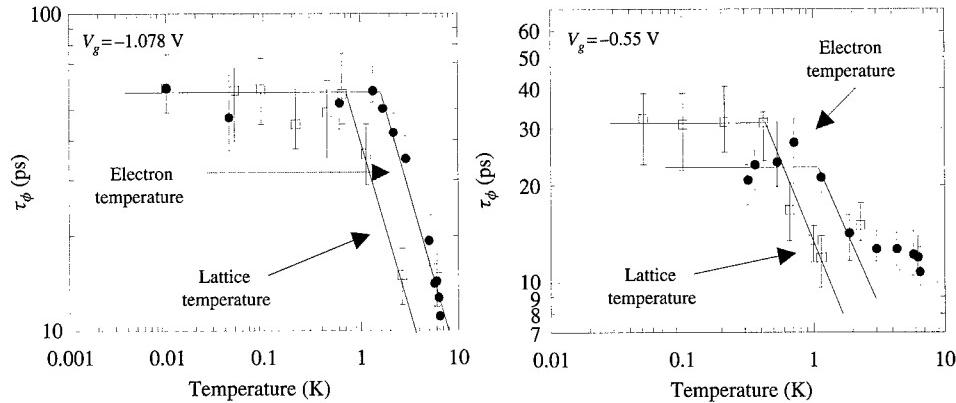
Fig. 2. Phase-breaking time as a function of lattice and electron temperature with $1/T$ fits.
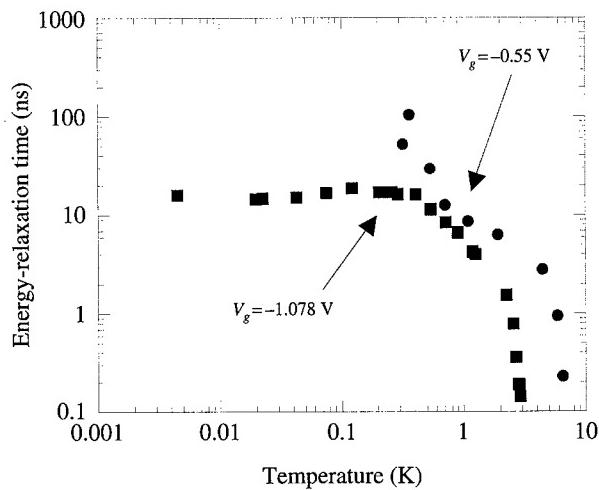

Fig. 3. Energy relaxation time as a function of the electron temperature.

electron scattering at higher gate voltages. At lower gate voltages, there seems to be an error in the fitting at high current values. This might be attributed to the fact that, at low gate voltages, a quantum wire behavior might be more dominant at higher current values, which leads to a $T^{-2/3}$ decay of the phase-breaking time. The observed energy relaxation time, as shown in Fig. 3, is much higher than the phase-breaking time, being on the order of a few tens of nanoseconds. The power per electron curve shows a break point at about 2.5 K, which signifies the onset of a regime where the decay of the energy relaxation time takes place. This would clearly suggest that the energy relaxation process is not an electron–electron scattering process.

Thus, we see that the phase-breaking time shows saturation at low temperature and currents and a T^{-1} dependence at higher temperatures, which is expected in lower dimensional systems with electron–electron scattering. The energy relaxation time is three orders of magnitude higher than the observed phase-breaking time and can be attributed to scattering by acoustic phonon modes. This scattering process is seen to set in at about 2.5 K and become the dominant one at higher temperatures. Future work needs to study the size

scaling effects of these parameters, by measuring open quantum dot arrays with equal numbers of dots of different sizes, and superlattice effects, by creating samples with single and multiple quantum dots.

References

[1] T. Ikoma, T. Odagiri, and K. Hirakawa, *Quantum Effect Physics, Electronics and Applications*, in *IOP Conf. Ser*, (1992).

[2] J. P. Bird *et al.*, Surf. Sci. **361/362**, 730 (1996).

[3] P. A. Lee, A. D. Stone, and H. Fukuyama, Phys. Rev. **B35**, 1039 (1987).

[4] J. P. Bird *et al.*, Phys. Rev. **B52**, 8295 (1995).

[5] J. P. Bird *et al.*, Phys. Rev. **B51**, 18037 (1995).

[6] D. Pivin *et al.*, Phys. Rev. Lett. **82**, 4687 (1999).

[7] C. M. Marcus *et al.*, Phys. Rev. **B48**, 2460 (1993).

[8] R. M. Clarke *et al.*, Phys. Rev. **B52**, 2656 (1995).

[9] D. K. Ferry and S. M. Goodnick, *Transport in Nanostructures* (Cambridge University Press, Cambridge, UK, 1997) Sec. 7.1.4.

On the use of Bohm trajectories for interpreting quantum flows in quantum dot structures

J. R. BARKER[†]

Nanoelectronics Research Centre, Department of Electronics and Electrical Engineering, The University of Glasgow, Glasgow G12 8LT, U.K.

R. AKIS, D. K. FERRY

Center for Solid State Electronics Research, Center for Systems Science and Engineering, and Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-5706, U.S.A.

Bohm trajectories derived from the probability density and probability current density reveal a simple picture of particle flow in quantum dots. A simple prescription is given showing how quantum vortices originate from the crossings of the underlying classical ray paths.

© 2000 Academic Press

Key words: quantum dots, quantum vortices, quantum transport.

1. Introduction

Recent simulation studies of magnetotransport fluctuations in open square quantum dots (see Fig. 1) have demonstrated the importance of the quantization of modes in the leads which causes the electrons to enter the dot in collimated beams [1–3]. This collimation has been conjectured to be responsible for a highly selective excitation of dot states. In addition, 'scarred wavefunctions' [3,4] have been observed in the simulations which are interpreted as having an amplitude which is maximized along a particular classical orbit. There is an obvious difficulty in interpreting the calculated probability densities as 'collimated flows' and 'classical orbits' because the orthodox interpretation of quantum mechanics precludes the existence of objective particle trajectories. Here, however, we re-examine the simulations [3] in the light of the trajectory methodology introduced by Bohm [5] which identifies the momentum of an electron at a point r as the spatial gradient of the phase of the wavefunction at that point. The velocity field $v(r, t)$ is thus very easy to calculate from the probability density $\rho(r, t)$ and the current density $j(r, t)$

$$v(r, t) = \frac{j(r, t)}{\rho(r, t)}. \tag{1}$$

Bohm's theory is mathematically identical to orthodox quantum mechanics and yields the same predictions [10], although it requires care when used as an *a priori* methodology in transport studies [6–9]. The advantage of the Bohm picture is that it permits the extraction of unambiguous velocity fields and spatial trajectories for electrons injected into a quantum dot without taking any classical limit.

[†]E-mail: J.Barker@elec.gla.ac.uk

Single aperture throttle   Double aperture open
   square dot
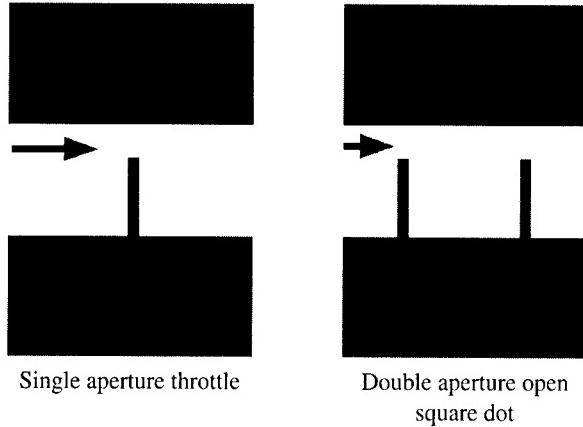
Fig. 1. Schematics of generic open quantum dot devices: single aperture throttle and the double aperture square dot.

2. Trajectories and streamlines

Previous studies have utilized numerical solutions to the steady state Schrödinger equation for square dots (0.3 μm^2) using a self-consistent soft wall model [3]. Here we shall use analytical techniques on hard wall models to extract the underlying physics. The Bohm picture projects the Schrödinger equation for $\Psi(x, y) = \sqrt{\rho}\exp[iS/\hbar]$ into two real equations: the continuity equation and (for the steady state) the energy conservation equation:

$$\nabla \cdot j = \nabla \cdot (\rho v) = 0 \tag{2}$$

$$E = T + V + V_Q = T + V - \frac{\hbar^2}{2m}\frac{\nabla^2\sqrt{\rho}}{\sqrt{\rho}} \tag{3}$$

where V_Q is the quantum potential, T is the classical kinetic energy $1/2mv^2$, where the velocity is related to the gradient of the phase by $v = (\nabla S - eA)/m$ (allowing for a magnetic field described by vector potential A). V is the quantum dot confinement potential.

The steady flow corresponding to (2) and (3) is described by two autonomous coupled equations for the particle position and velocity:

$$\frac{dr}{dt} = v = (\nabla S - eA)/m \tag{4}$$

$$\frac{dv}{dt} = \{-\nabla(V + V_Q) + ev \times B\}/m. \tag{5}$$

The autonomous system (4) and (5) describes velocity streamlines or trajectories which cannot cross because the wavefunction is single-valued [8–10]. The topology of the velocity flow may be established by using stability analysis to characterize the flow in the vicinity of the fixed points.

3. Singularities in the flow

There are two classes of singularity in the flow. Linear stability analysis around the velocity nodes, points r_0 for which $v = 0$, reveals that for zero magnetic field the local flow pattern is hyperbolic; for nonvanishing magnetic field there occur in addition closed elliptical orbits due to the Lorentz force. The hyperbolic fixed points (saddle points) act to separate the flow, whereas the closed orbits surrounding fixed points which are

centres correspond to Lorentz vortices. A second class of vortex flow occurs even in the absence of a magnetic field when $\nabla \times \mathbf{v} = 0$ everywhere except at the nodes in the amplitude of the wavefunction [8, 9]. At the amplitude nodes the phase and hence the velocity is not defined. This is a purely quantum effect. The closed orbital motion around the centre of the quantum vortex is due to the quantum potential V_Q (see eqn (3)), which in the vicinity of points where $\sqrt{\rho} = 0$, may exhibit a $|r - r_0|^{-2}$ dependence leading to closed orbits. In the case of the 2D quantum dot structures of Fig. 1 we therefore expect three classes of flow: (a) open orbits which connect the entry and exit leads of the dot; (b) hyperbolic orbits which link to the saddle points or terminate at hard wall boundaries; (c) Lorentz and quantum vortices. In the following we shall ignore the Lorentz vortices by setting $\mathbf{B} = 0$.

4. Single aperture

The key elements of the velocity flows in the open square dot may be discovered in a simple analysis of the single aperture throttle (Fig. 1). Consider an incident propagating wave in a mode given by the wavevector $\mathbf{k} = (k_n, n\pi/W)$ in the upstream lead region, where $k_n = [k^2 - (m\pi/W)^2]^{\frac{1}{2}}$ is the longitudinal component and $n\pi/W$ $(n = 1, 2, \ldots)$ is the transverse component. For clarity we set $\hbar = m = 1$. In an obvious notation, the solution upstream and downstream of the aperture (of width a at $x = 0$) is given by:

$$\Psi_{out}^U(x, y; n) = e^{ik_n x}\phi_n(y) + \sum_{m=1}^{M} S_{mn}^U e^{-ik_m x}\phi_m(y) \tag{6}$$

$$\Psi_{out}^D(x, y; n) = \sum_{m=1}^{M} S_{mn}^D e^{ik_m x}\phi_m(y) \tag{7}$$

where the $\{\phi_m(y) = (2/W)^{\frac{1}{2}}\sin(m\pi y/W); 0 < y < W\}$ are the transverse components of the wavefunctions. $\{S_{mn}^{U,D}\}$ are the S-matrix elements. At the aperture itself the solution is a superposition of aperture modes (y-axis points downwards)

$$\Psi_{in}^A(x = 0, y) = \sum_{p} C_p \sin(p\pi y/a) \qquad 0 < y < a = 0 \text{ otherwise.} \tag{8}$$

The S-matrix elements and the components C_p may be obtained by matching using the orthonormality of the transverse modes and the boundary conditions; Dirichlet boundary conditions at the hard walls; continuity of the wavefunction and its x-derivative at the aperture.

For large incident transverse and longitudinal wavevectors the transmission coefficient is approximately a/W, the classical value. Under these conditions the coefficients C_p are only significant for mode numbers $p \approx na/W$. The S-matrix S_{mn} elements are similarly concentrated around $m = n$. Under these conditions and defining $k_x = k_n$; $k_y = n\pi/W$; $\theta = \arctan(k_y/k_x)$, we find the downstream scattered wave in the form:

$$\Psi_{out}^D \approx \frac{1}{2}\exp[i(k_x + k_y\tan\theta - \pi/2)]\{\Phi(y - x\tan\theta) - \Phi(y + x\tan\theta)\}. \tag{9}$$

Here the periodic function $\Phi(y)$ is defined by:

$$\Phi(y) = \sum_{p} C_p \exp(ip\pi y/a) \qquad \text{for } |y - 2Na| < a$$

$$\Phi(y) = 0 \qquad \text{for } |y - 2Na| > a \tag{10}$$

where the integer N is defined such that $(y - 2Na)$ lies in the domain $[-W, W]$. At $x = 0$ the outgoing wave coincides with the aperture wavefunction eqn (8) as required. For $x > 0$ the outgoing wave is focused within a reflecting ray path forming a strip of vertical width a and having straight sides inclined at
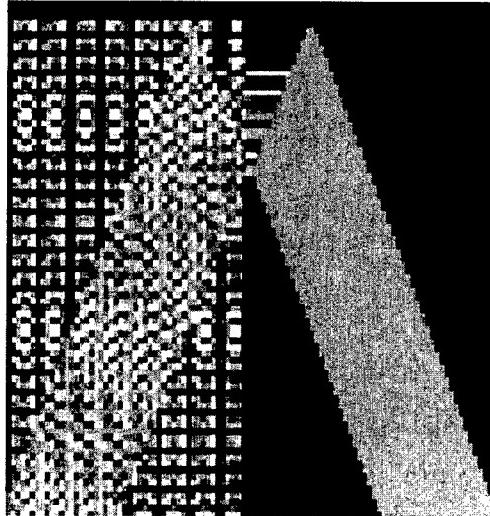
Fig. 2. Upstream and downstream flows in throttle device.

angles $\pm\theta$ to the x-axis. On the upward and downward paths the dominant part of the wave approximates to $\exp[i(k_x x \pm i k_y y)]$. In these regions the Bohm velocity components are simply $v_x = \hbar k_x/m$; $v_y = \pm\hbar k_y/m$. The aperture thus acts to filter out the propagating components $\exp[\pm i k_y y]$ from the incident transverse mode. The boundaries of the ray paths are determined by classical ray trajectories reflecting off the walls at angles $\pm\theta$. Within the ray path the downstream wave simply consists of the wave $\exp[i(k_x x - i k_y y)]$ and its reflections in the boundary walls including a π phase change for each reflection (consistent with hard wall boundary conditions). Near the boundaries the upward moving component interferes with the downward component to select only the longitudinal motion where $\Psi \approx \sin(k_y y) \exp[i k_x x]$. The exact solution reveals diffraction effects and small amplitude modulation superposed on the simple downstream wave. The velocity streamlines undulate and gradually spread away from the classical ray paths as demonstrated in an earlier analysis [11].

The upstream wave may be shown exactly to form a superposition of a standing wave $\sin(k_y y)\sin(k_x x)$ (formed from the incident wave and its reflection at $x = 0$) and the mirror image of the downstream wave $\Psi^D_{out}(-x, y)$. The upstream probability density field is therefore a standing wave pattern scarred by interference with the reflecting ray paths. On a coarse scale the Bohm velocity field is therefore very similar to the downstream pattern. However, in fine scale the interference with the standing wave pattern generates vortex chains. The open streamlines thread their way through the vortex chains into the aperture and finally join the simple downstream flow. This simple model is illustrated in Figs 2 and 3.

5. Double aperture

A similar analysis of the square dot, essentially a double aperture device, reveals similar features to the throttle. Standing wave patterns are formed upstream of the first aperture and within the square dot itself. At the first aperture a strong propagating wave is injected which is approximately confined to a ray path of width determined by the aperture and which involves multiple reflections of the boundary walls before reaching the exit aperture. The reflecting ray paths interfere with the standing wave in the dot to form micro-vortices as with the upstream case in the single aperture. In addition, new and larger vortices form where the ray path loops back on itself due to reflections at the walls. At high wavevectors this is the dominant feature. The
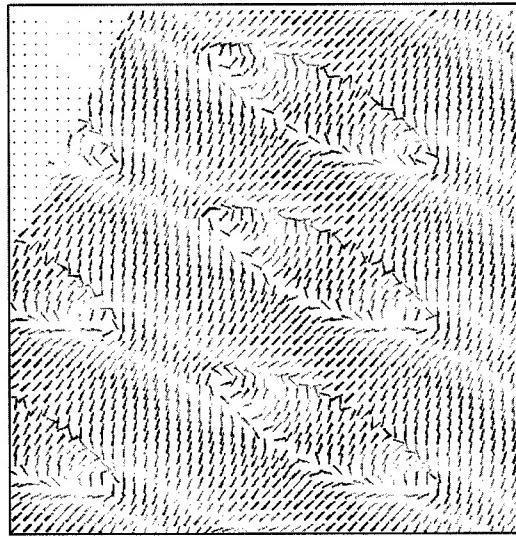
Fig. 3. Detail of streamlines in the vicinity of scarring by the classical orbit interfering with the standing wave states showing looping between vortices.
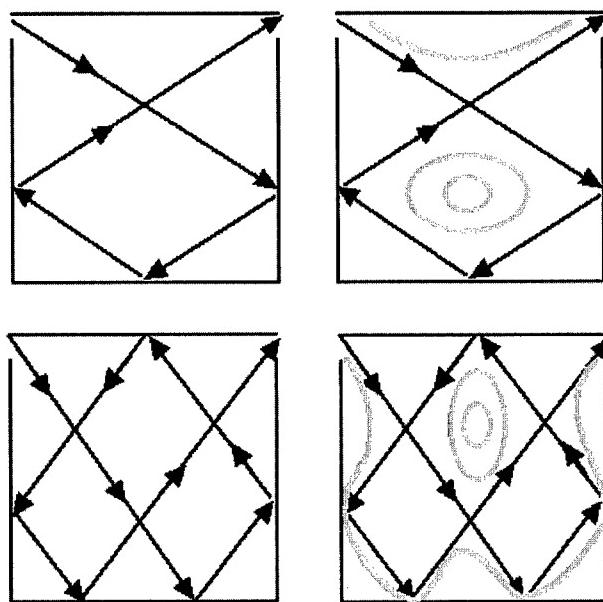
Fig. 4. Construction of vortices and unbounded flow from classical bounce orbits.

origin of the vortices may be seen in Fig. 4. To construct the vortices we imagine a classical path formed by a free particle bouncing between the walls of the dot where the injection direction is determined from the incident wave-vector components. Because quantum trajectories cannot cross, the quantum case is derivable from the classical cases by letting the classical streamlines repel each other near crossing points. The resulting
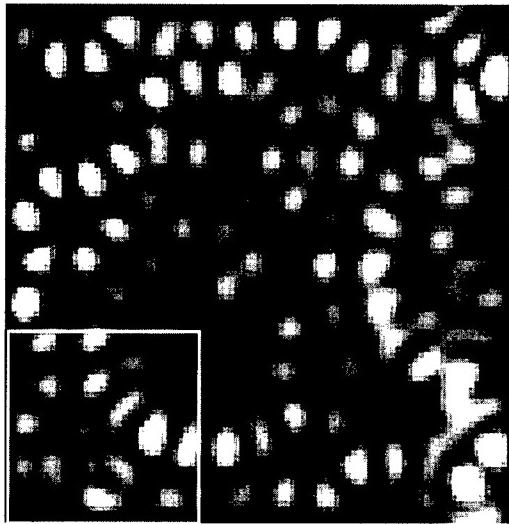
Fig. 5. Numerical calculation for probability density in a square dot using data from [3].
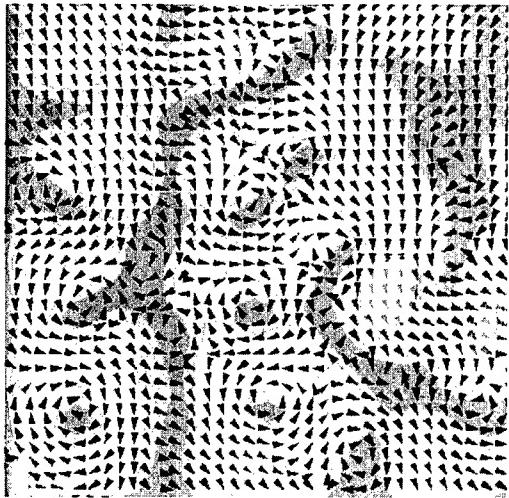

Fig. 6. Velocity vector field for inset region of Fig. 4.

flow topology contains hyperbolic and vortex flows as shown in Fig. 4. This result is confirmed by analytical and numerical calculations. The net effect is that the complex classical ray trajectories associated with the quantum ray propagation separates the flow into localized flow around vortex centres and meandering open orbits which thread between them. Typically there are several open orbit paths some with minimum traversal times, others with very long paths and hence long traversal times. It is likely that dissipative processes will disrupt the latter flows. Beyond the exit aperture we again find a simple ray path propagation. These general features, although derived from a simple model, show clearly in our further analysis of data published in [3]. An example of the probability density field for a 0.3 μm square dot is shown in Fig. 5. Figure 6 shows clearly both vortices and hyperbolic flow separation in the computed velocity field for the region shown inset

in Fig. 5. The complete flow field for Fig. 5 corresponds qualitatively to the simple flow shown in Fig. 4 plus the additional flow complexity following from interference with the standing waves in the scarred regions.

6. Conclusions

It is concluded that the previously conjectured collimated states in open quantum dots correspond to highly focused meandering Bohm trajectories which link the square dot arena. It may also be shown that the presence of a magnetic field produces noticeable curvature in the trajectories as might be expected from the action of the Lorentz force. There is no precise analogy between Bohm trajectories and the equivalent classical orbits because of vortex formation which follows from the requirement that quantum orbits cannot cross. Nevertheless, a large portion of the flow topology and vortex formation can be understood from the underlying classical orbits and their crossing point topology. There are scars in the probability density close to propagating (nonclosed) velocity streamlines but the flow is interspersed with vortex pairs which may considerably distort the flow from any classical counterpart.

References

[1] J. P. Bird, R. Akis, D. K. Ferry, J. Cooper, K. Ishibashi, Y. Ochia, Y. Aoyagi, and T. Sugano, Semicond. Sci. Technol. **13**, A4 (1998).

[2] I. V. Zozoulenko, A. S. Sachrajda, P. Zawadzki, K.-F. Bergrren, Y. Feng, and Z. Wasilewski, Semicond. Sci. Technol. **13**, A7 (1998).

[3] R. Akis and D. K. Ferry, Semicond. Sci. Technol. **13**, A18 (1998).

[4] E. J. Heller, Phys. Rev. Lett. **53**, 1515 (1984).

[5] D. Bohm, Phys. Rev. **85**, 166 (1952).

[6] J. R. Barker, Semicond. Sci. Technol. **4**, 911 (1994).

[7] J. R. Barker, S. Roy, and S. Babiker, S., *Science and Technology of Mesoscopic Structures*, edited by S. Namba, C. Hamaguchi, and T. Ando (Springer-Verlag, London, 1992) Chap. 22, p. 213.

[8] J. R. Barker, Extended Abstracts, *1998 Sixth International Workshop on Computational Electronics* (IWCE-6) IEEE Catalog No. 98EX116,1–4, (1998).

[9] J. R. Barker and D. K. Ferry, Semicond. Sci. Technol. **13**, A135 (1998).

[10] P. R. Holland, *The Quantum Theory of Motion* (Cambridge University Press, Cambridge, 1993).

[11] Hua Wu and D. W. L. Sprung, Phys. Lett. **A196**, 229 (1994).

Current effects on high magnetic field resonances in a dot array

L.-H. LIN, N. AOKI, K. NAKAO, Y. OCHIAI

Department of Materials Science, Chiba University, 1-33 Yayoi, Inage, Chiba 236-8522, Japan

F. GE, A. ANDRESEN, C. PRASAD, J. P. BIRD, D. K. FERRY

Department of Electrical Engineering and Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85287-5706, U.S.A.

K. ISHIBASHI, Y. AOYAGI, T. SUGANO

Nanoelectronic Materials Laboratory, Frontier Research Program, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

In recent studies of disordered quantum dot arrays, it was found that a re-entrant metal–insulator transition occurs in these structures as the gate voltage is varied, even though there is no change in carrier density or disorder. Here, we study resonances that appear with the quantum Hall effect in these dot arrays. Several resonances are observed on the side of the plateaus, which may be due either to charging of isolated islands or to transmission resonances in the quantum point contacts. Heating of the carriers causes these peaks to decrease with an energy relaxation time comparable to that of the dots themselves.

© 2000 Academic Press

Key words: quantum dots, magnetotransport, heterostructures.

In recent months, studies of small arrays of open quantum dots have revealed new and interesting behavior, with onset of exponential localization at low temperatures, and a novel localization–delocalization transition at still lower temperatures [1, 2]. It is found that carrier heating in these structures occurs with a characteristic energy-relaxation time [3] that is significantly different from the phase-breaking time [4]. In this paper, we report studies on resonant peaks that appear in the zeros of the longitudinal resistance in the quantum Hall effect regime. These resonances can be either resonant reflection (in the longitudinal resistance) or resonant transmission (in the longitudinal conductance), but are thought to be coherent features of the overall transmission in the dot array [5]. We study the temperature and current dependence of these peaks, and show that they are characterized by a relaxation time that is closer to the energy-relaxation time rather than the phase-coherence time, a result that seems at odds with their coherent nature. The basic device structure is the same as that discussed in [1–4], although this array has only three dots, which are formed by bias applied to standard 'split' gates on the surface of the material.

In Fig. 1A, we plot the resistance as a function of temperature for this array. The resistance shows a sharp rise for T below about 2 K, and then a slower, almost logarithmic behavior for $T < 0.5$ K. This overall behavior characterizes a localization with the resistance diverging as the temperature approaches zero, and
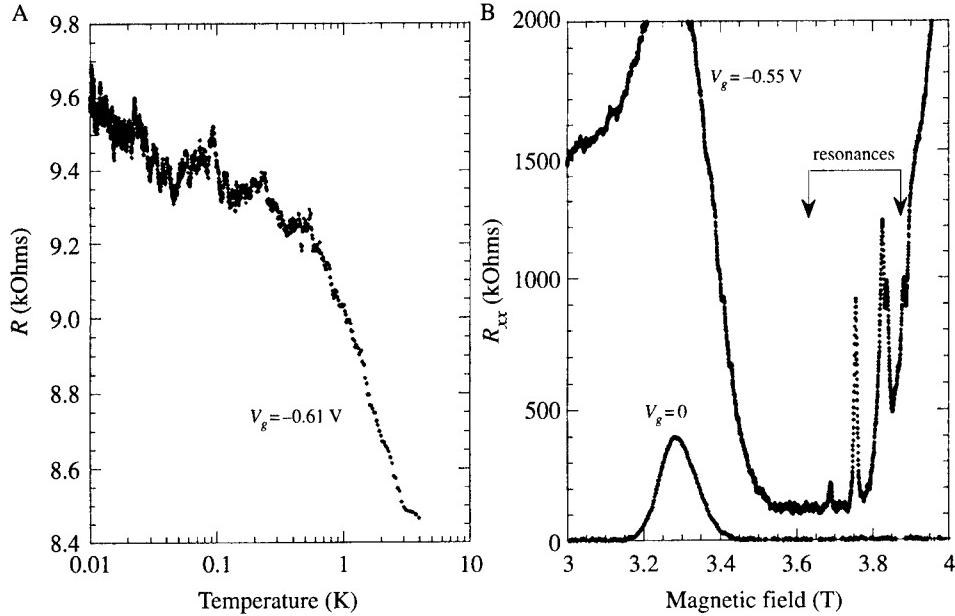
Fig. 1. A, Resistance of the dot array as the temperature is varied. B, Comparison of the longitudinal magnetoresistance of the dot array and the bare electron gas. Our interest is in the resonances indicated by the arrows. The curve for $V_g = -0.55$ V has been offset vertically by 100 Ohms.

is not present in the absence of bias on the depletion gates. Note that there is no weak localization, and the absence of the resistance increase at zero bias suggests that the material is not disordered to any extent. In Fig. 1B, we compare R_{xx} in the dot array with that of the bare two-dimensional electron gas. It is apparent that dot confinement leads to a narrowing of the plateau as well as several resonances. It is the behavior of these resonances that is of interest here. To our knowledge, these resonances are not limited to a specific plateau.

In Fig. 2, we illustrate the changes in these resonances that occur as the temperature is raised (panel A) and as the current through the device is raised (panel B). From these figures, we can see that increasing bias current leads to carrier heating that has a comparable effect to lattice heating. At the temperatures of interest here, the primary scattering mechanism is carrier–carrier scattering and some residual random potential scattering (from impurities and defects). In both cases, the single temperature in the scattering process is that of the electrons themselves, so electron heating should display a comparable effect to lattice heating. We note, however, that the current dependence of these peaks is much stronger than that found for the resistance itself at low magnetic field. We also note that the resistance peak below the plateau is considerably enhanced by the formation of the dot, which is an additional localization of this peak, even though the Fermi level is presumably within the bulk Landau level.

From the temperature dependence of the resonant peaks, we can establish an electron temperature for each value of bias current. That is, at a given bias current, the peak height is measured and compared with that of the temperature dependence. This determines the electron temperature. The power input to the array is related to the dissipation process through,

$$VI = Nk_B(T_e - T_0)/\tau_r,$$

where N is the total number of electrons in the dot array [3]. From this relationship, we can then determine the relaxation time and its dependence upon the electron temperature. This is shown in Fig. 3. It appears that
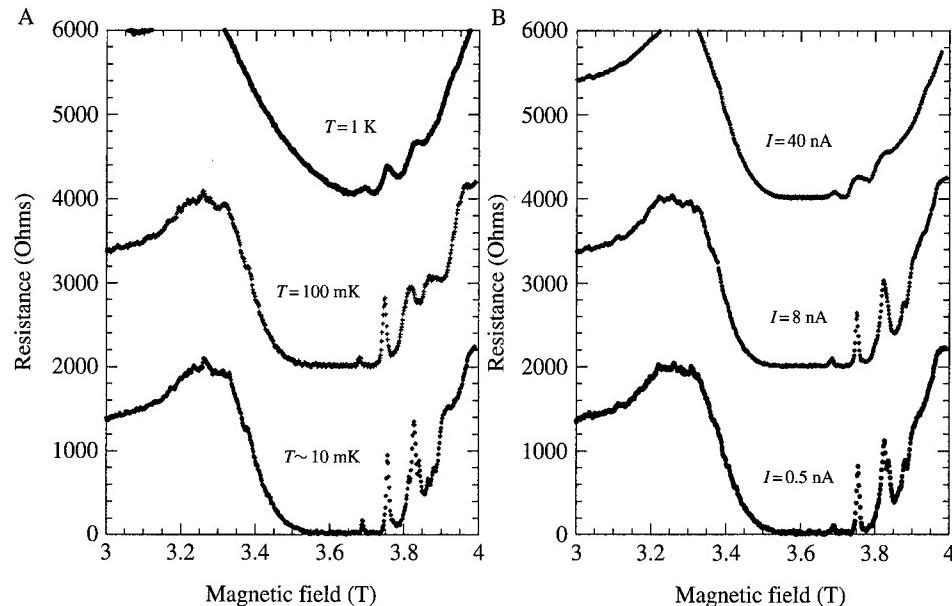
Fig. 2. The longitudinal resistance for the dot array with $V_g = -0.55$ V. A, R_{xx} for several values of the lattice temperature. The upper two curves have been offset by 2000 and 4000 Ohms, respectively, for clarity. B, R_{xx} for several values of the current through the array. The upper two curves have been offset by 2000 and 4000 Ohms, respectively, for clarity.
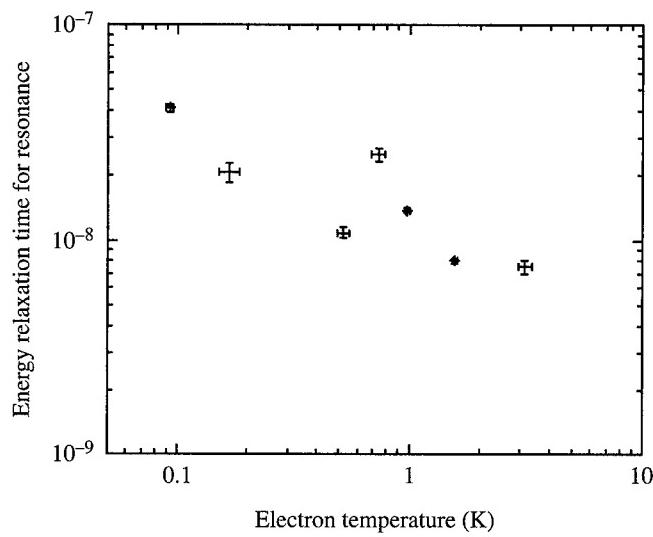
Fig. 3. The relaxation time (in seconds) determined from the dependence of the electron temperature on the bias current.

there is an overall decay with electron temperature, although it is expected to actually be more or less constant below about 1 K, from studies of the zero magnetic field behavior [3]. The magnitude of the relaxation time found here is quite close to the energy-relaxation time determined by a similar process at zero magnetic field. While there is certainly some uncertainty in the number of carriers contributing to the resonant peaks (one

might assume that only the upper spin level of the 3/2 Landau level would contribute to such a resonance, although we have used the total number of carriers), the observed phase-breaking time is orders of magnitude smaller [4]. Consequently, it appears that these coherent peaks are destroyed by a normal energy relaxation process rather than a simple phase-breaking process.

In conclusion, we find that coherent resonances in the longitudinal resistance of a quantum dot array have an electron heating dependence that seems to be determined by the energy-relaxation time rather than the phase-coherence time.

References

[1] A. Shailos *et al.*, Superlatt. Microstruct. **27**, 311 (2000).
[2] A. Andresen *et al.*, Phys. Rev. **B60**, 16050 (1999).
[3] L.-H. Lin *et al.*, Phys. Rev. **B60**, R16299 (1999).
[4] C. Prasad *et al.*, Superlatt. Microstruct. **27**, 315 (2000).
[5] P. J. Price, Microelectron. J. **30**, 925 (1999).

Eigenstates in quantum dots confined by an inhomogeneous magnetic field

R. AKIS, D. K. FERRY

Center for Solid State Electronics Research & Department of Electrical Engineering, Arizona State University, Tempe AZ 85287-5706, U.S.A.

We study the eigenstates in quantum dots in which electrons are confined by the application of an inhomogeneous perpendicular magnetic field, focusing on the effect that the specific details of the shape of confining field has on determining these states. In contrast to the edge state picture established in studies on circular dots, we find that dots with more irregular geometries show a more complicated behavior in the interior of the dot. In particular, we find that certain states show indications of having their amplitude enhanced along particular classical periodic orbits in the interior, a phenomenon known as 'scarring'.

© 2000 Academic Press

Key words: quantum dots, magnetic confinement, periodic orbits.

In recent years, there has been growing interest in mesoscopic structures in which electrons are confined by the application of inhomogeneous perpendicular magnetic fields rather than by electrostatic potentials. This can be achieved experimentally, for example, by positioning a superconducting disk over a GaAs heterostructure [1]. In studies of two-dimensional electron gases (2DEGs) in which a nonuniform magnetic field was applied, calculations showed that a step in the magnetic field would create an effective potential [2]. Magnetic barriers have the unusual property that their relative height is dependent on the wave number of the incident electrons. The differences in the penetration depth can be understood classically in terms of how much of a cyclotron orbit that an electron can complete before being transmitted or reflected, with larger angles of incidence leading to greater penetration [2]. Besides quasi-one-dimensional barrier structures, magnetic quantum dots have also been studied theoretically, assuming a *circular* confining geometry [3, 4, 5]. In this case, the resulting vector potential that enters the Hamiltonian shares the circular symmetry and angular momentum, m, remains a good quantum number. For finite m, the magnitude of the typical eigenstate shows a series of one or more concentric rings, leading to their interpretation as being edge states guided along the magnetic field boundary.

In this paper, we consider the effect that changing the boundary shape of the magnetic dot has on the eigenstates. As we shall demonstrate, when the dot no longer has circular symmetry, the basic edge state picture no longer suffices and rather complicated behavior can occur in the interior of the dot. Moreover, we find that certain states show indications of being 'scarred', that is, their amplitude appears to be enhanced along particular classical periodic orbits that pass through the interior.

Our first example is a semicircular magnetic dot of base diameter a, with the magnetic field profile shown
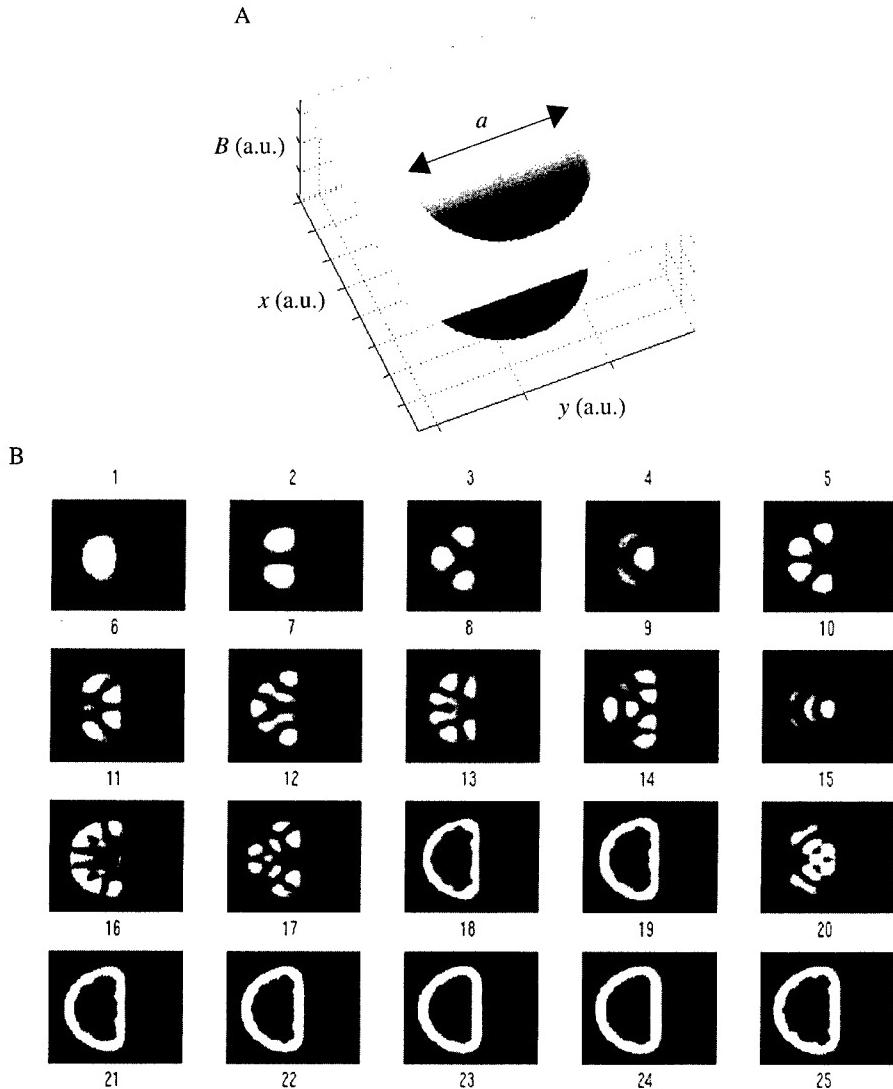
Fig. 1. A, Geometry of one of the inhomogeneous magnetic confining fields used in our simulations, here forming a semicircular quantum dot. B, The first 25 eigenstates of the semicircular dot, given $B = 307 B_a$. What is plotted is the *magnitude* of the wavefunction versus x and y.

in Fig. 1A. As with all the dots we study, the magnetic field, B, is taken to evolve essentially as a step, with $B = 0$ in the interior and finite outside. To find the eigenstates, we solve this 2D problem on a finite difference mesh, with the magnetic field accounted for through the introduction of appropriate Peierl's phases on the off-diagonal terms of the Hamiltonian. Since the resulting matrix is sparse, we use the ARPACK implementation of the Arnoldi factorization method to solve the problem numerically [6].

Figure 1B shows the first 25 eigenstates of the semicircular dot, with $B = 307 B_a$ (B_a is the field for which the magnetic length is equal to a). In choosing this geometry, we have combined the round boundary of the circle with the straight boundary of a simple barrier. In view of the latter, it is important to emphasize that since the barrier is *magnetic*, it is not a source of simple mirror reflections. That symmetry is broken.
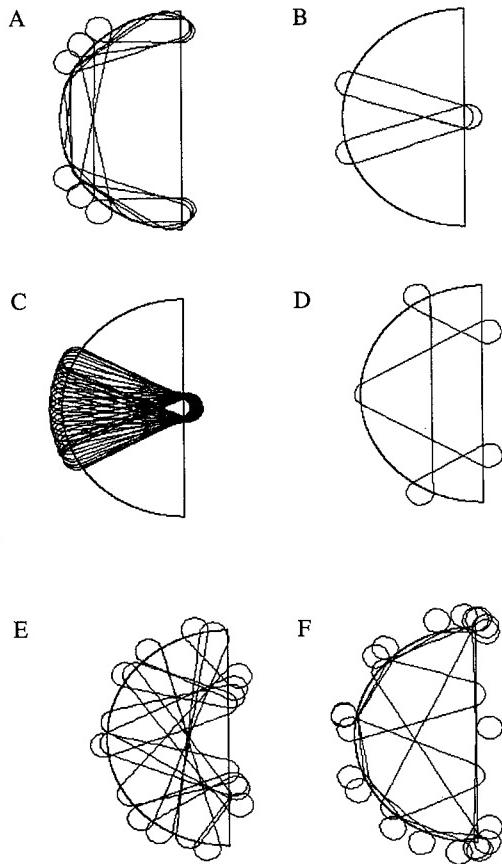
Fig. 2. Here, A–F correspond to six different orbits supported by this structure, given a cyclotron radius of $\sim$0.05a.

Several of the eigenstates shown here obviously correspond to edge states that travel along the outer edge of the *entire* boundary. However, some of the states show rather interesting behavior in the interior. For example, states 7 and 15 have a distinct 'V' form, while state 10 looks somewhat like a 'fan'. State 11 has edge state-like properties, however, its amplitude is concentrated mainly along the circular boundary. State 12 has an 'A' pattern to it, while states 21 and 22 show very complicated interior patterns that create a low amplitude spot on the center of the right side. State 20 shows a similar pattern apparently coexisting with an edge state.

To get some insight into the physics of this structure, we have also performed classical simulations, finding many of the periodic orbits or closed trajectories allowed by its geometry. Several examples are shown in Fig. 2. Figure 2A shows a 'whispering gallery'-type orbit that follows the left circular boundary while simply bouncing off the right barrier. In Fig. 2B, we show an example of a 'V'-type orbit. In Fig. 2C is a more complicated but closely related orbit (obtained using slightly different initial conditions) that is in the form of a 'fan'. Figure 2D shows an orbit that has an 'A' pattern. Figure 2E shows a rather complicated orbit that does not visit a spot on the center of the right boundary. The orbit in Fig. 2E has a similar 'blind' spot, but spends much of its time on a path that follows the entire perimeter of the dot. There is a resemblance between the patterns revealed by these orbits and several of the eigenstates pointed above. We believe that many of these eigenstates are in fact 'scarred' by these orbits, or at least orbits that belong to the same or similar
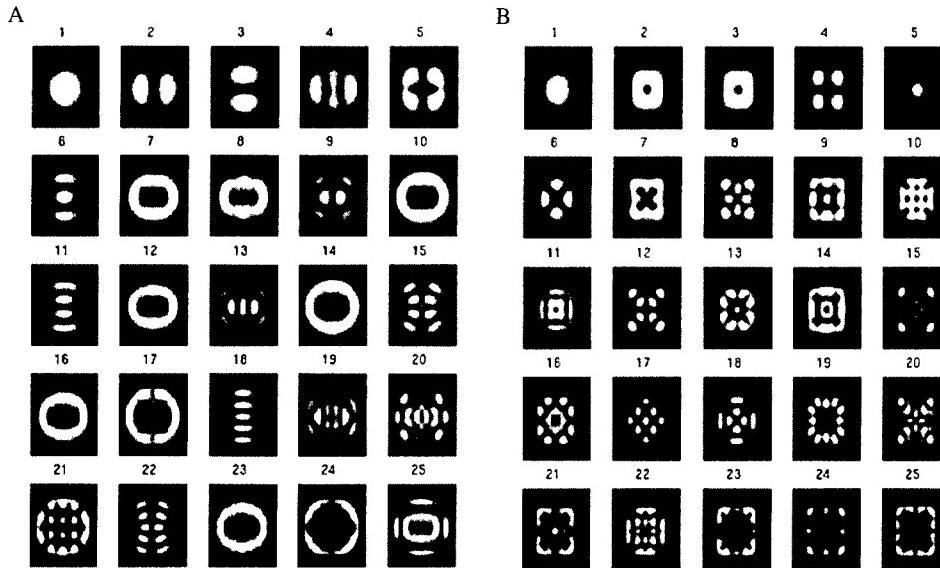
Fig. 3. A, The first 25 eigenstates of the stadium, with a middle section $a/6$ units long. Here, $B = 350\ B_a$. B, The first 25 eigenstates of a square with $B = 303\ B_a$, a is now the side of the square.

'families'. As Heller *et al.* [7] have pointed out, it can be quite difficult to make an exact correspondence between a specific orbit and 'scar', since orbits tend to have similar looking 'cousins', which differ in the number of reflections they have against the boundaries. The 'V' and the 'fan' orbits are an example of this phenomenon, with the latter being a more complicated variation of the same basic pattern. As one might expect, Poincaré plots show that these orbits are in close proximity in phase space. A detailed examination of the classical phase space is in progress.

Analogous results occur in other geometries. In Fig. 3A, we show eigenstates for a stadium-type geometry. The stadium differs from the circle in that a rectangular section is included to separate the two semicircular halves. As above, there is a mix of 'edge-like' and interior filled states. States 6, 11, and 18 clearly reflect the presence of what are known as 'bouncing ball' orbits, which occur when an electron is multiply reflected between parallel barriers. In Fig. 3B are eigenstates for a square geometry. Standing out here are states with patterns forming an 'X' (7, 12, 15, 20 and 21) and a '+' (8, 10, and 18). We have found classical orbits conforming to these sorts of patterns, as well as more complex ones such as state 17, occurring in this geometry.

In closing, we must note that relationships between classical orbits and the eigenstates of the circular dot have been established [3, 4, 5]. However, because of the continuous radial symmetry, a given orbit is degenerate with an infinite number of equivalent orbits that have different angular orientations, leading to the ring structure alluded to above. In the dots studied here, that symmetry has been is broken, which is the reason why orbits that correspond to a *particular* orientation can be preferentially selected and 'scar' the eigenstates.

References

[1] A. K. Geim, I. V. Grigorieva, S. V. Dubonos, J. G. S. Lok, J. C. Maan, A. E. Filippov, and F. M. Peeters, Nature **390**, 259 (1997).

[2] F. M. Peeters and A. Matulis, Phys. Rev. **B48**, 15166 (1993).

[3] L. Solimany and B. Kramer, Solid State Commun. **96**, 471 (1995).

[4] H.-S. Sim, K.-H. Ahn, K. J. Chang, G. Ihm, N. Kim, and S. J. Lee, Phys. Rev. Lett. **80**, 1501 (1998).

[5] J. Reijneirs, F. M. Peeters, and A. Matulis, Phys. Rev. **B59**, 2817 (1999).

[6] Website: `http://www.caam.rice.edu/software/ARPACK/index.html`.

[7] E. J. Heller, P. W. O' Connor, and J. Gehlen, Phys. Scr. **40**, 354 (1989).

Temperature-dependent conductance in single and multiple quantum dots: the role of the disorder

F. BADRIEH, R. AKIS, D. K. FERRY

Center for Solid State Electronics Research and Department of Electrical Engineering, Arizona State University, Tempe AZ 85287-5706, U.S.A.

We present calculations of conductance, G, as a function of inverse temperature, $1/T$, for a quantum dot array with disorder, a clean quantum dot enclosed by soft boundaries and a simple two resonance model. The behavior in each case is qualitatively similar, with G appearing to increase exponentially provided one starts at a minimum. Unlike some recent experimental studies, we do not see any logarithmic behavior at low temperatures.

© 2000 Academic Press

Key words: quantum dots, disorder, transport.

Recent experimental measurements of the conductance in single and multiple quantum dots, with varying amounts of disorder, have yielded interesting temperature-dependent behavior [1, 2]. In analogy to the localization exhibited by two-dimensional systems [3], it was found that the temperature (T) dependent variation of the conductance (G) may be fitted as

$$G = G_0 + G_1 \ln(T) + G_2 \exp\left[-\left(\frac{T_0}{T}\right)^p\right].$$ (1)

The G_0 term is thought to represent the quantum point contact (QPC) contribution to the overall conductance. The logarithmic term was found to dominate at low temperatures ($T < 0.5$ K) and is also observed in studies of the two-dimensional metal–insulator transition in MOSFETs [3] The physical origin of this term remains unclear at present, although it may be related to an electron–electron interaction effect [4].

In this paper, we try to account for the T-dependent behavior seen in experiment by using two different simulations methods: (1) the recursive Green's function (RGF) approach [5] and (2) a stabilized variant of the transfer matrix (SVTM) approach [6]. In both cases, G is determined using the Landauer formula. Finite T is taken into account by convoluting G with the derivative of the Fermi function.

Figure 1A shows the geometry for a dot array simulated using RGFs. In this case, disorder is represented by a random potential on each lattice site with maximum amplitude W, as per the standard Anderson model. As one may expect, the results shown in Fig. 1B indicate that reducing W and increasing the QPC width s results in increased G. It also appears to result in a reduced number of resonant peaks in G. In Fig. 1C, we now consider the case of a single quantum dot with *no* disorder and a soft confining potential (see [7] for more details) modeled using the SVTM method. This structure is much smaller than the array, so the result is fewer resonances over a comparable range of Fermi energies.
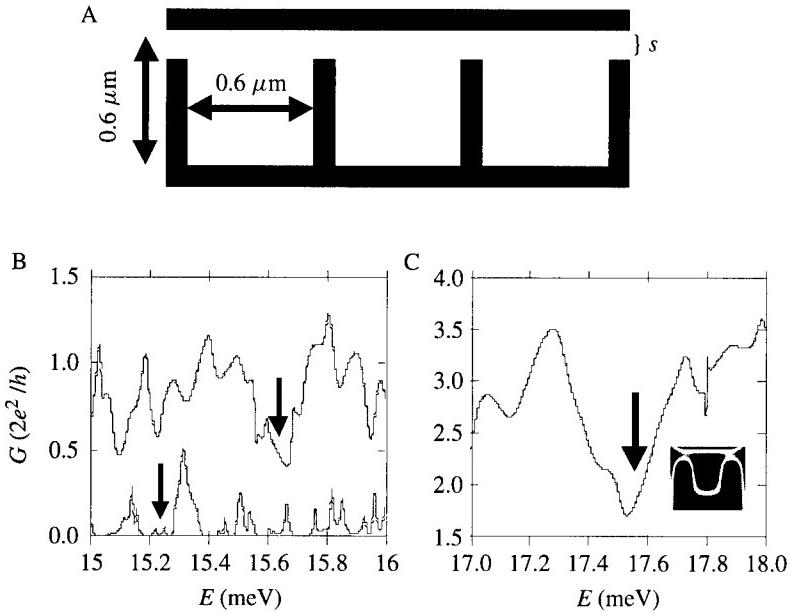
Fig. 1. A, Geometry of the split-gate quantum dot array that is simulated. B, G of the array versus Fermi energy, E, for $s = 0.1$ μm, $W = 18$ meV (lower curve) and $s = 0.15$ μm, $W = 4$ meV (upper curve). C, G versus E for a 0.4 μm dot with soft boundaries as depicted in the inset.

In Fig. 2, we examine the evolution of G as a function of $1/T$ for the three examples considered. The arrows in Fig. 1 indicate the starting point that has been chosen, positioned somewhere in a local minimum in G in each case. Each curve shown here shows the same general behavior—starting relatively flat, then gradually increasing until passing over a maximum. The maximum arises when the tail of the Fermi derivative is of sufficient amplitude that the adjacent peaks in G get picked up in the convolution. The dashed lines are fits obtained using eqn (1), but leaving out the $\ln(T)$ term. The exponential term does a relatively good job of fitting G while it is increasing. Comparing the fitting parameters given in the figure caption, one notes that T_0 is 0.13 K in the higher disorder but more pinched off quantum dot array, significantly smaller than in the other two cases. We again note that this case shows a higher density of resonances. The T_0 parameter has been thought to reflect an energy gap, which here might be interpreted in terms of the distance between surviving resonances (as a structure is made more open, more resonances become smeared out). Another important factor in determining T_0 is where one is situated between two resonances.

To confirm our hypothesis that the behavior shown here is simply a resonance effect, in Fig. 3A we show the results of an *ad hoc* double resonance model where the Fermi energy is fixed in between the two resonances (see inset). The dashed line is the exponential fit (G_0 has been subtracted out here). Particularly evident here is that the fit falls below the curve at large $1/T$. In Fig. 3B the solid line is the fit obtained to a set of actual experimental data [1, 2] using the *full* eqn (1). The dashed line is the result obtained by scaling the double resonance curve and adding the same G_0 and $G_1 \ln(T)$ terms used in the experimental fit. Note that this curve is above the fitted line at high $1/T$ and falls below it at low $1/T$, *which is what the experimental data actually does*.

It is important to note that the $\ln(T)$ term had to added to the simulation results. The simple thermal smearing we use here does not produce this T dependence under any circumstances, regardless of dot geometry or disorder (we have used several different models all having the same basic effect). This suggests to us that additional physical effects must be included to account for this behavior, with electron–electron
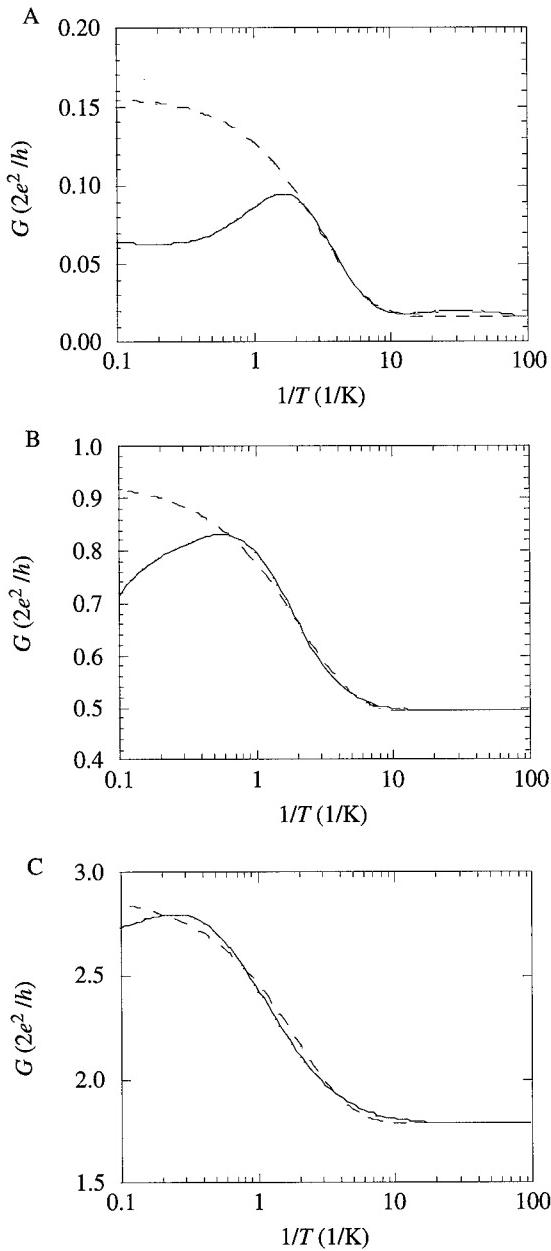
Fig. 2. A, G versus $1/T$ for the array with $E_F = 15.245$ meV (solid line). The dashed line is a fit to eqn (1) with $G_0 = 0.02\ e^2/h$, $G_1 = 0$, $G_2 = 0.28\ e^2/h$, $T_0 = 0.13$ K, and $p = 1.25$. B, G versus $1/T$ for the array with $E_F = 15.633$ meV (solid line). The dashed line is a fit using $G_0 = 0.986\ e^2/h$, $G_1 = 0$, $G_2 = 0.88\ e^2/h$, $T_0 = 0.47$ K, and $p = 1.2$. C, G versus $1/T$ for the soft dot with $E_F = 17.55$ meV (solid line). The dashed line is a fit using $G_0 = 3.576\ e^2/h$, $G_1 = 0$, $G_2 = 2.2\ e^2/h$, $T_0 = 0.52$ K, and $p = 1.1$.

interactions being a likely candidate, as mentioned above. It should also be noted that thermal smearing makes G tend towards an average value as a function of E, with the $T = 0$ minima and maxima lying below and above the finite T curve. In contrast, the experimental data always shows an upward shift at
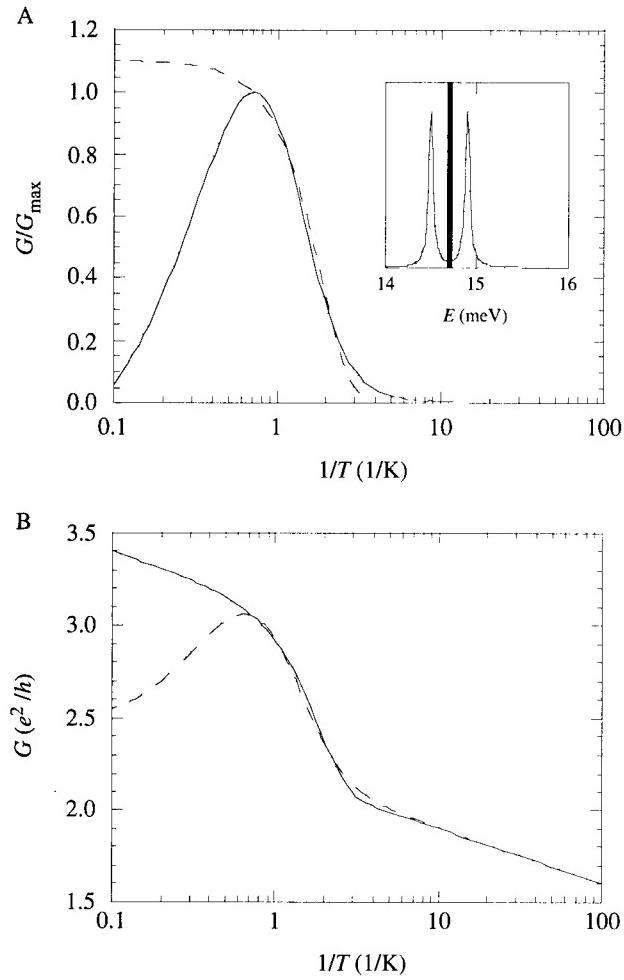
Fig. 3. A, Normalized conductance versus $1/T$ using the two resonance model (solid line). The dashed line is a fit to eqn (1) with $G_0 = G_1 = 0$, $G_2 = 1.1$, $T_0 = 0.54$ K, and $p = 2.4$. Inset: the two transmission resonances, with the line in between indicating $E_F = 15.7$ eV. B, G versus $1/T$ using eqn (1) with $G_0 = 2.2\, e^2/h$, $G_1 = 0.13\, e^2/h$, $G_2 = 0.90\, e^2/h$, $T_0 = 0.54$ K, and $p = 2.4$. The dashed line was obtained by scaling the curve in A by a factor of 0.81 and adding $G_0 = 2.2\, e^2/h$ as well as a $\ln(T)$ term with prefactor $G_1 = 0.13\, e^2/h$.

finite T, with very few $T = 0$ maxima lying above the finite T curve. Once again, this behavior is only reproduced in the simulations by adding a $\ln(T)$ term. It should be emphasized that the curves obtained in Fig. 2 were all obtained by starting somewhere in a relative G minimum. In general, the finite T behavior is more complicated and depends sensitively on the starting point. Moving away from the starting energies we have chosen in our examples, one can obtain curves with additional structure such as extra points of inflection or oscillations superimposed on top. The experimental curves do not appear to have the same degree of sensitivity to the starting point, and the exponential behavior is more generic. That said, it is worth noting that if an appropriately scaled, $\ln(T)$ is added to a curve that has more complex structure, we find that it can overwhelm some of these details and still generate a curve not unlike that already shown in Fig. 3B.

In conclusion, we believe that the exponential behavior observed experimentally might be accounted for, at least in part, by the resonance effect noted here, while the logarithmic behavior requires additional physics to be introduced.

References

[1] A. Andresen *et al.*, Phys. Rev. **B60** (Rapid Communications), 16050 (1999).

[2] L.-H. Lin *et al.*, Phys. Rev. **B60**, R16299 (1999).

[3] For recent overviews of the metal-insulator transition in two dimensions, see: E. Abrahams, Physica **3**, 69 (1998); V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, *ibid.* **3**, 79 (1998).

[4] B. L. Altshuler *et al.*, Phys. Rev. Lett. **44**, 1288 (1980).

[5] H. U. Baranger, D. P. DiVincenzo, R. A. Jalabert, and A. D. Stone, Phys. Rev. **B44**, 10637 (1991).

[6] T. Usuki, M. Saito, M. Takatsu, R. A. Kiehl, and N. Yokoyama, Phys. Rev. **B52**, 8244 (1995).

[7] J. P. Bird, R. Akis, and D. K. Ferry, Phys. Rev. **B60**, 13676 (1999).

Effect of boundary conditions on quantum-anti-dots in the presence of a magnetic field

F. BADRIEH, D. K. FERRY

Department of Electrical Engineering and Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85287-5706, U.S.A.

Many people have studied the conductance properties through an array of anti-dots, especially since the observation of Weiss oscillations. In most cases, however, in which the recursive Greens functions are used on a spatial lattice, periodic boundary conditions are employed. In this paper, we analyse the effects of boundary conditions and magnetic field on the conductance behavior in a number of anti-dot-shaped, GaAs/AlGaAs 2DEG quantum systems. The effect of periodic boundary conditions causes a reduction in the overall conductance. The effect of changing the boundary conditions is more profound for lower numbers of anti-dots.

© 2000 Academic Press

Key words: quantum anti-dots, magnetotransport, heterostructures.

Over the past several years, there has been considerable interest in anti-dot arrays, in which scattering potentials are sited in an otherwise metallic quasi-two-dimensional electron gas. Interest was stimulated by the observation of resistance peaks correlated to commensurate orbits around one or a few anti-dots [1, 2]. Many theoretical studies of these additional peaks, termed *Weiss oscillations*, have focused upon the chaotic nature of transport in the array when the magnetic length is incommensurate with the period of the array [3, 4, 5]. These studies seemed to indicate that large transverse currents would be induced at the resistance peaks, an effect which is precluded by the absence of any transverse *terminal* currents in the Hall configuration. Usually, the studies considered one or a few anti-dots with periodic boundary conditions, which do not represent the imposition of normal Hall configuration constraints. In this paper, we analyse the effects of boundary conditions and magnetic field on the conductance behavior in a number of anti-dot arrays in the GaAs/AlGaAs two-dimensional electron gas. Five different anti-dot arrays were considered, and two runs were done for each to compute the conductance versus magnetic field: the first with transverse hard-wall boundary conditions and the second with periodic transverse boundary conditions. The size of the anti-dot region in each case was 1×1 μm and we used the recursive Green's functions method in the Landauer formalism. The comparisons are made among 1, 3, 5, 7 and 9 anti-dot systems where the anti-dot potential shape follows a sinusoid raised to the fourth power. In all cases, the conductance is plotted versus magnetic field in the range $0 < B < 1$ T. The overall effect of the magnetic field is to depopulate the energy levels in the system and hence decrease the number of propagating modes. This in turn causes a decrease in the conductance curves. In the absence of the anti-dot potential, the conductance curves assume a step-like
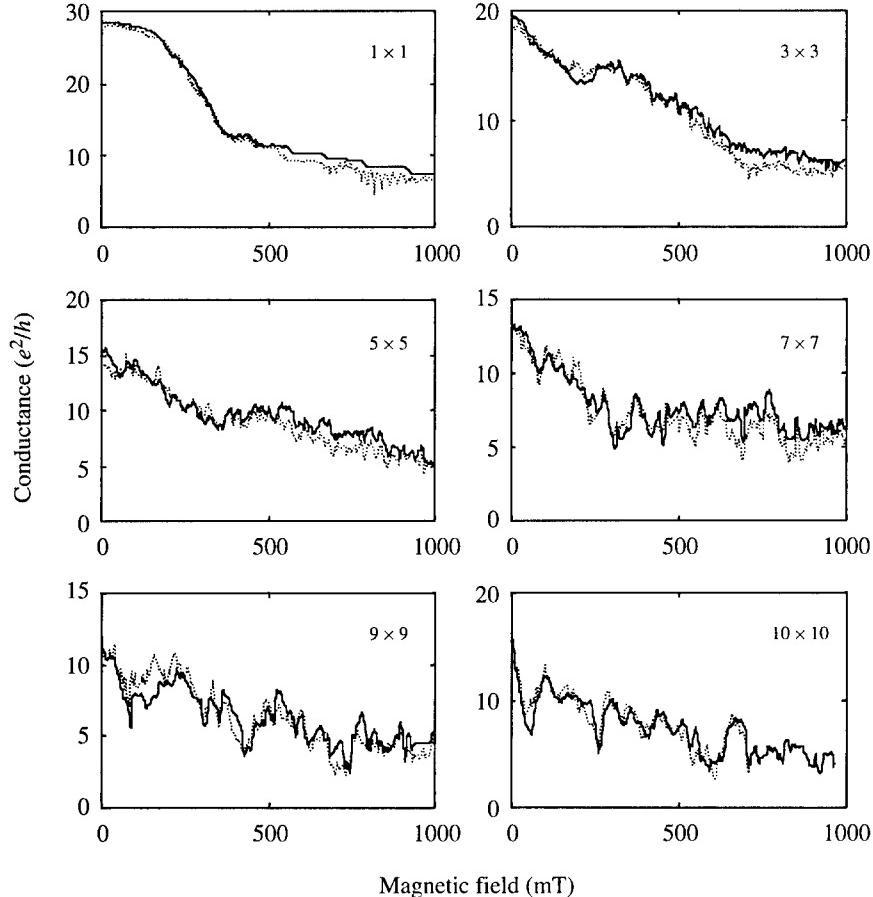
Fig. 1. The conductance through the anti-dot array as a function of the magnetic field normal to the two-dimensional plane of the 1 μm^2 square active area. The number of anti-dots in the active area is indicated by the label of each panel. The solid curves are for hardwall boundary conditions, while the dotted curves are for periodic boundary conditions.

shape, which decreases with magnetic field. The steps are of equal length when plotted versus $1/B$. With the introduction of the anti-dot potential, the step-like behavior begins to wash out, although the conductance reduction behavior persists, and an oscillatory behavior emerges. Increasing the number of anti-dots causes a reduction in the maximum of the conductance curves, and tends to reduce the magneto-resistance. The effect of periodic boundary conditions causes a reduction in the overall conductance while maintaining the decreasing behavior. The effect of changing the boundary conditions is more profound for smaller numbers of anti-dots.

The anti-dot potential is applied to each site in the discretization used in the recursive Green's function approach [6]. Here, the potential is

$$V(x, y) = V_0 \sin^4(i\pi x/L) \sin^4(i\pi y/W)$$

where W and L are the width and length of the active region, and i is the number of anti-dots in each direction, taken to be the same. The choice of exponent is arbitrary, but this choice gives sufficiently hardwall effects from the anti-dots. The choice of hardwall or periodic boundary conditions appears in the unperturbed on-site Green's function itself, and these are modified for the particular case. In addition, a small random potential is
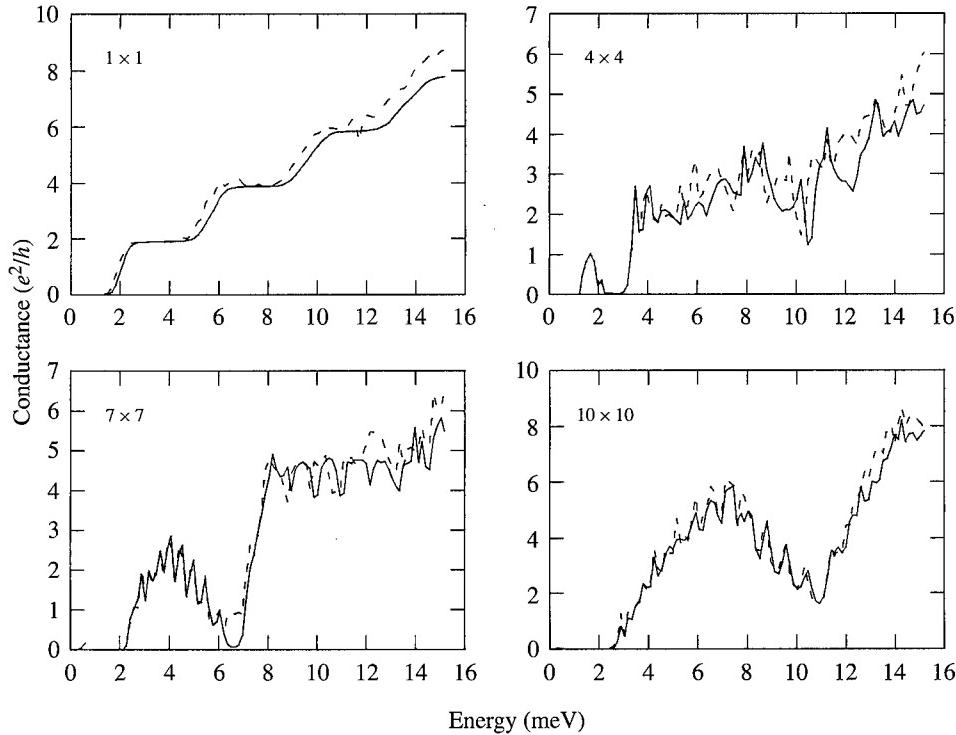
Fig. 2. Variation of the conductance as the Fermi energy (density) is varied in the anti-dot array.

added to the system to represent inherent disorder that exists. In Fig. 1, we plot the conductance as a function of the magnetic field for a 1×1 μm active area. The Fermi energy is set to 14.5 meV, corresponding to a carrier density of approximately 4×10^{11} cm^{-3}, and V_0 is some four times larger. The different panels are for $1 \times 1, 3 \times 3, 5 \times 5, 7 \times 7, 9 \times 9, 10 \times 10$ anti-dot arrays. It is clear that the number of anti-dots affects the overall transmission through the structure, primarily as a result of narrowing of the open channels between the anti-dots themselves and between the boundary and the anti-dots in the hardwall case. This conductance, at vanishing magnetic field begins near 30 (in units of e^2/h) and decreases by a factor of 3 as the number of anti-dots is increased. This is mainly due to the fact that there are about 30 propagating modes in the single anti-dot case, and this is reduced to roughly 10–15 propagating modes in the case of the 10×10 array. We note, however, that most of this decrease occurs in the low magnetic field region, so that the overall magnetoresistance is also decreased as the number of anti-dots increases. Indeed, the number of modes for $B = 1$ T decreases from roughly 20 modes to about 8 in these same two cases. The cases for hardwall boundary conditions are shown as solid lines, while the cases for periodic boundary conditions are shown by the dotted lines. The presence of hardwall boundary conditions does not significantly affect the resulting conductance curves, although the overall conductance is *larger* in this case by a small amount. However, none of the peaks or dips in conductance of Fig. 1 correspond to expected positions of the Weiss resonances.

We have also studied the behavior as the Fermi energy is varied, and one sees typical conductance steps as the latter is raised and modes are turned on in the structure, as shown in Fig. 2. As the number of anti-dots is increased, this step structure is broken up by the interferences of paths encircling the anti-dots. By increasing numbers of anti-dots, the turn-on of conductance is also increased, although there is little effect of the difference in the two boundary conditions imposed.

The results of these simulations seem to indicate that the conductance through the anti-dot array, to a large degree, is insensitive to the boundary conditions that are applied to the simulation. This would indicate that the conductance is probably not going to show large transverse currents as has been suggested by studies of chaotic dynamics in such arrays. Increasing the number of anti-dots, in a fixed simulation area, however, does seem to affect the conductance through the array, which is probably due to the reduction in transverse width of any single open path through the array. In addition, there are quasi-resonances which lead to large localized increases in the density within the array, probably due to standing wave interferences between the two ends of the current path. These appear particularly prominent in studies in which the Fermi energy (density) is varied within the structure. However, none of the prominent peaks or dips in conductance correlate well with the expected positions of the Weiss resonances.

References

[1] R. Gerhardts, D. Weiss, and K. von Klitzing, Phys. Rev. Lett. **62**, 1173 (1989).
[2] D. Weiss, *Localization and Confinement of Electrons*, edited by F. Kuchar, H. Heinrich, and G. Bauer (Springer-Verlag, Berlin, 1990) pp. 247–257.
[3] T. Geisel, J. Wagenhuber, P. Niebauer, and G. Obermair, Phys. Rev. Lett. **64**, 1581 (1990).
[4] R. Fleischmann, T. Geisel, and R. Ketzmerick, Phys. Rev. Lett. **68**, 1367 (1992).
[5] T. Ando, S. Uryu, and S. Ishizaka, Jpn. J. Appl. Phys. **38B**, 308 (1999).
[6] D. K. Ferry and S. M. Goodnick, *Transport in Nanostructures* (Cambridge University Press, Cambridge, UK, 1997).

Three-dimensional modelling of dissipative quantum transport in quantum dots and atomistic scale devices using nonHermitian generalized potentials

J. R. BARKER[†], J. R. WATLING

Nanoelectronics Research Centre, Department of Electronics and Electrical Engineering, University of Glasgow, Glasgow G12 8LT, U.K.

In this paper we introduce a phenomenology for inserting dissipation into the single-particle Schrödinger equation for carrier transport by utilizing appropriate nonHermitian additions to the Hamiltonian. The nonHermitian terms are determined by incorporating model particle trapping/de-trapping, momentum gain/loss, energy gain/loss into the quantum continuity equations derived within the Bohm picture and then reconstructing the full Hamiltonian by reversing the Bohm projection. The new phenomenology is designed to obtain quantum velocity flows using the Bohm projection of solutions to the nonHermitian Schrödinger equation for applications in 2D and 3D quantum dots and mesoscopic MOSFETs. For this purpose we introduce a novel fast algorithm to compute the wave function in 2D and 3D based on a two time step iteration and direct integration.

© 2000 Academic Press

Key words: quantum transport, Bohm picture, quantum dots, mesoscopic.

1. Introduction

Quantum transport in quantum dot structures at low temperatures and *a fortiori* quantum transport in room-temperature ultra-small MOSFET devices in the atomistic limit require consideration of dissipative processes. The latter include: carrier trapping/de-trapping and phase breaking processes corresponding to momentum, angular momentum and energy loss processes in the electronic configuration space due to carrier–carrier, carrier–phonon and carrier–impurity centre scattering processes. There is a broad divide between the classical relaxive Boltzmann transport model used to describe MOSFET devices and the full quantum many-body formalism for treating open quantum dots. There is a clear need for a linking formalism that preserves the simplicity of the particle or classical hydrodynamic approach used in device modelling and yet incorporates the essential features of quantum mechanics. There is an opportunity to do this in the mesoscopic regime where the transport is partly coherent and where there is occupancy of a large number of channel modes (large wave vectors) due to high-temperature operation. Equivalently, from the classical standpoint, the carrier trajectories begin to deviate from classical flow due to size–wavelength commensurability. In a recent study of square quantum dots [1] we have shown that the quantum coherent velocity flow is carried by open orbits which meander between localized current voices (Fig. 1). It is clear that although the short

[†]E-mail: J.Barker@elec.gla.ac.uk
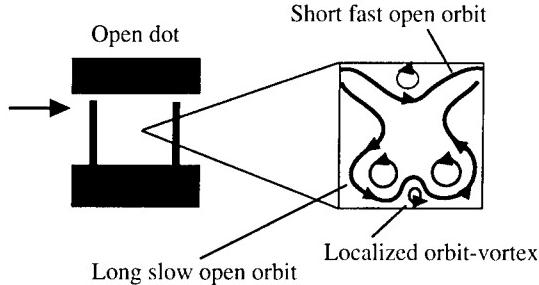
Fig. 1. The main velocity flow processes in a square quantum dot include open orbits which directly traverse the dot, and long meandering open orbits which thread through clusters of localized vortex flows [1].

open orbits may be traversed within the coherence time this is not likely for the vortex flows or the longer open orbits. The simple picture will thus break down due to dissipative processes. Here we shall show that the simple velocity flow picture developed from the Bohm picture of quantum mechanics [2, 3, 4, 5, 6, 7] can be extended to include dissipation.

Our long-term objective is to seek simple quantum corrections to particle- and hydrodynamic-based device models [7]. The quantum potential cannot be used here unless it has been derived from the wave function because it cannot account for quantum vortex distortions of the flow which depend on the nodal topology of the wave functions [7], nor is it clear how it is altered by dissipation.

2. The Bohm picture: Hermitian Hamiltonian

In recent studies it has been demonstrated [4, 5, 6, 7] that the Bohm picture of quantum mechanics [2, 3] establishes a clear correspondence between classical velocity fields and quantum velocity flows when there is large number of occupied modes—the mesoscopic regime—provided the transport is coherent and ballistic. In the Bohm picture it is possible to derive the continuity equations for classical-like dynamical variables such as the particle probability density $\rho(r, t)$, the velocity field $v(r, t)$, the energy density and so on, by direct projection from the Schrödinger equation for the Hermitian Hamiltonian $H = p^2/2m + V(r)$ (this simple form is chosen without loss of generality):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot j = 0 \tag{1}$$

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + v \cdot \nabla v = -\nabla(V + V_Q)/m \tag{2}$$

$$\frac{dr}{dt} = v = \nabla S/m. \tag{3}$$

Here the quantum potential $V_Q(r, t) = -(\hbar^2/2m)(\nabla^2 \sqrt{\rho})/\sqrt{\rho}$ and the velocity $v(r, t) = \nabla S(r, t)/m = j(r, t)/\rho(r, t)$ are derived from the amplitude $\sqrt{\rho}$ and phase S of the wave function (they are secondary variables). The Bohm picture envisages a carrier to exist on one of the trajectories of the flow defined by (1) and (2) according to the probability density and with a velocity given by the gradient of the local phase S. Because the wave function is single-valued, the quantum trajectories $r(t)$ defined by (2) and (3) cannot cross; in particular the velocity streamlines cannot cross. The velocity flow topology is consequently determined by the fixed points of (2) and (3); they are either hyperbolic flows about saddle points at the nodes of the velocity field or vortex flows about centres which are the nodes of the wave function.

3. The Bohm picture: nonHermitian Hamiltonian

A new phenomenological quantum mechanical formalism has been introduced by us [8] which incorporates generalized scalar and vector nonHermitian potentials which lead to a simple description of dissipative quantum transport within a single-body formalism based either on the carrier wave function or density matrix. Here we examine how to construct such nonHermitian Hamiltonians with the aid of the Bohm picture. Out strategy is to insert classical phenomenological energy-loss, momentum-loss, particle trapping/de-trapping directly into the corresponding Bohm picture continuity equations. Then by reconstructing the Schrödinger picture we recover the appropriate nonHermitian extensions to the Hamiltonian. The new Hamiltonian is then used to determine the wave function and the quantum velocity flows are then calculated from $v = j/\rho$.

For dissipation models the re-constructed Hamiltonian may be written as a nondissipative part H_0 plus a nonHermitian part iH_1 where H_1 is Hermitian and has the general structure:

$$H_1 = E_0 + H_a(p) + H_b(r) + H_c(r, p). \tag{4}$$

The constant term E_0 leads to a description of global carrier trapping/de-trapping depending upon its sign. The term H_b may be devised to produce localized source and sink terms. The term H_a is a function of the momentum operator and leads to global momentum and energy gain–loss processes; the term in H_c may be devised to introduce local gain–loss processes.

Two examples are presented. The first adds phenomenological generation and recombination terms to the continuity equation (1):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot j = G - R \tag{5}$$

$$G = \frac{\rho \sum_\mu d\delta(r - r_\mu)}{\tau_g}; \qquad R = -\frac{\rho \sum_\mu d\delta(r - r'_\mu)}{\tau_r}. \tag{6}$$

The equivalent Hamiltonian is reconstructed as the nonHermitian form:

$$H = H_0 + \frac{i\hbar}{\tau_g} \sum_\mu d\delta(r - r_\mu) - \frac{i\hbar}{\tau_r} \sum_\mu d\delta(r - r'_\mu) \tag{7}$$

which describes trapping/de-trapping at a discrete set of locations. The corresponding quantum velocity flow now contains new singularities to determine the flow topology. They are source and sink fixed points or attractor and repeller basins for continuous trapping/de-trapping regions. Figure 2 sketches a pair of trap absorption/trap emission centres in the velocity phase portrait. The streamlines correspond to possible trajectories for a particle to be removed from the flow at one attractor and re-emitted elsewhere from the repeller. Figure 2 is essentially a picture of scattering with capture/re-capture from the Bohm viewpoint.

A second example is the re-constructed nonHermitian Hamiltonian

$$H = (1 - i\beta[\theta(t - t_1) - \theta(t - t_2)] + i\beta[\theta(t - t_1 - \tau_E) - \theta(t - t_2 - \tau_E)])\frac{p^2}{2m} + V. \tag{8}$$

The nonHermitian term acts to remove kinetic energy in the time range (t_1, t_2) and restores kinetic energy in the time range $(t_1 + \tau, t_2 + \tau)$. This leads to frictional processes in the flow during the time interval t_1, $t_2 + \tau$: they derive from a velocity-dependent quantum friction term and a viscosity component in the velocity flow equation. The particle continuity equation also acquires sink and source terms acting at different times. The corresponding velocity flows again contain pairs of source and sink singularities which scatter particles within the flow at different energies.

4. Consequences for velocity flows in devices

The dissipative terms introduce new singularities into the velocity flows which, for example, cause particles to move off the closed orbits of the vortices (the vortices may also degenerate into spiral attractors
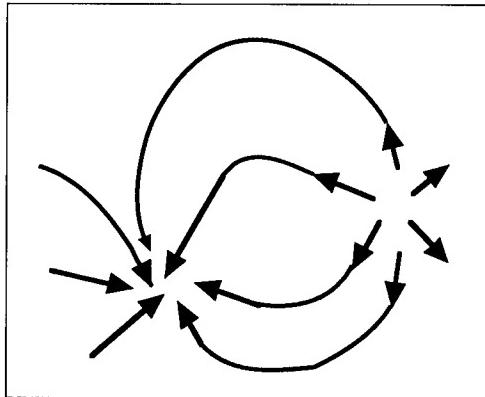
Fig. 2. New singularities in the velocity flow resulting from the nonHermitian Hamiltonian model. The case shown corresponds to a source and sink pair which model the scattering of a particle with consequent energy and momentum gain or loss.

re-emerging from source singularities to re-enter the flow either on a new vortex or an open trajectory). The situation has analogies to transport in quantising magnetic fields; where scattering-induced hopping between Landau centres induces dissipative current flow. Similarly, particles on long traversal time orbits may be removed and re-inserted into the flow following scattering events.

5. Application to 2D and 3D wave packet transport in small devices

The dissipative quantum transport formalism described briefly here is being used to study dissipative propagation in model 2D and 3D problems, including the dissipative transport of injected wave packets in throttles and quantum dots, and the flow of wave packets close to a rough interface within a MOSFET channel. The wave functions are derived from the Schrödinger equation with a nonHermitian Hamiltonian, and the velocity flows are extracted by computing the ratio of the current density to the probability density. The simulations are based on a new, direct method for solving the finite-differenced time-dependent 2D or 3D nonHermitian Schrödinger equation which avoids the huge computational complexity of the implicit Crank–Nicholson scheme (routinely used for 1D problems but prohibitive in higher dimensions).

The iterative scheme is briefly a two-time step direct method which unlike the Euler method has good convergence and stability:

$$\psi^{n+1} = \psi^{n-1} + \Lambda \cdot \psi^n \tag{9}$$

where n denotes the time step, ψ denotes a p-point structure (vector in 1D, matrix in 2D) derived from the discretization of the wavefunction over p-dimensional space; Λ denotes a banded $p–D$ matrix. The algorithm is simply derived from the evolution equations:

$$\Psi(r, t \pm \tau) = \exp[\mp i H \tau] \Psi(r, t)$$
$$\Psi(r, t + \tau) = \Psi(r, t - \tau) + \{\exp[-i H \tau / \hbar] - \exp[i H \tau / \hbar]\} \Psi(r, t) \tag{10}$$
$$\Psi(r, t + \tau) \approx \Psi(r, t - \tau) + \{-2i H \tau / \hbar\} \Psi(r, t)$$

in the limit of a small time step τ.

This method requires no matrix inversion and for 3D problems has only a complexity level of order $30N^3$. It is stable and convergent for similar criteria to the Crank–Nicholson scheme.

Figure 3 illustrates the method by showing frames from the time-evolution of a longitudinal Gaussian wave packet injected into a transverse mode of a quantum throttle device (compare [1]). The plot shows a 3D perspective of the colour-coded probability density as a function of position in the 2D structure. The exit flow shows strong forward streaming due the occupancy of several longitudinal wavevectors.
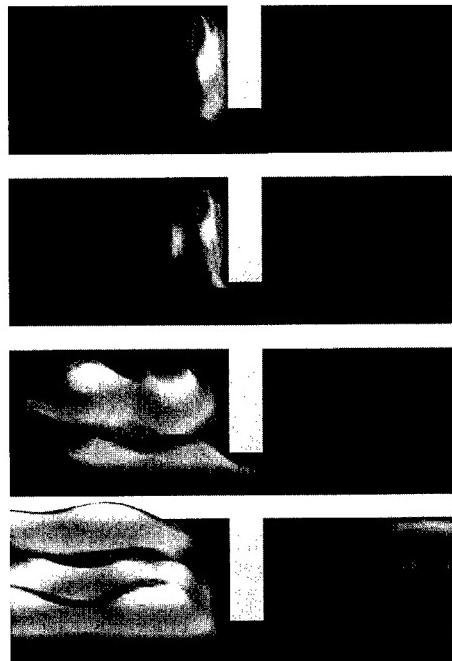
Fig. 3. Frames from a simulation of the time evolution of a Gaussian wave packet injected into a single transverse mode of a throttle device in 2D showing the probability density landscape as a function of time using the fast algorithm described in the text.

6. Conclusions

This paper has demonstrated a simple phenomenological approach to dissipative quantum transport derived using the Bohm trajectory picture augmented by phenomenological dissipative terms in the continuity equations. The equivalent Hamiltonian is found to be nonHermitian and solutions of the resulting Schrödinger equation will describe dissipative processes. Alternatively, the same solutions may be used to uncover the unexplored regime of dissipative quantum hydrodynamic flow which now includes sources and sinks for the velocity streamlines in addition to the normal vortex and hyperbolic flows.

References

[1] J. R. Barker, R. Akis, and D. K. Ferry, these proceedings.

[2] J. R. Barker and D. K. Ferry, Semicond. Sci. Tech. **13**, A135 (1998).

[3] D. Bohm, Phys. Rev. **85**, 166 (1952).

[4] P. R. Holland, *The Quantum Theory of Motion* (Cambridge University Press, Cambridge, 1993).

[5] J. R. Barker, Semicon. Sci. Technol. **4**, 911 (1994).

[6] J. R. Barker, S. Roy, and S. Babiker, in *Science and Technology of Mesoscopic Structures*, edited by S. Namba, C. Hamaguchi, and T. Ando (Springer Verlag, London, 1992) Chap. 22, p. 213.

[7] J. R. Barker, Extended Abstracts, *1998 Sixth International Workshop on Computational Electronics* (IWCE-6) IEEE Catalog No. 98EX116, 1–4 (1998).

[8] D. K. Ferry and J. R. Barker, Appl. Phys. Lett. **74**, 582 (1999).

Optically controlled confined potential of a vertically aligned double-dot array system

K. FUJII, T. YOSHIZAWA, T. OHYAMA, K. OTO, S. TAKAOKA, K. MURASE

Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan

K. GAMO

*Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-0043, Japan*

In a vertically coupled dot, it is expected that generalized Kohn theorem is invalid. Then electron–electron interaction can be observed in far-infrared absorption. This interaction depends on the coupling between dots. We fabricated a vertically aligned double-dot array using a double heterostructure to investigate the coupling. The strength of the coupling depends on the confine potential of the dot. In far-infrared magneto-optical absorption measurements under pulsed photoexcitation, we observed peak shift of asymmetry states and increasing of absorption intensity of symmetry states of the dot. This is a result of the confined potential modulation by light. This indicates the possibility of optical control of the coupling.

© 2000 Academic Press

Key words: vertically coupled quantum dot, far-infrared absorption, potential modulation by light.

1. Introduction

For the last decade, electronic states in a semiconductor quantum dot array have been investigated intensively by far-infrared (FIR) absorption measurements. FIR magneto-optical absorption measurements are a useful tool for the investigation of electronic states in a quantum dot. In the dot structure, however, lateral confinement potential is treated as a parabolic [1]. The generalized Kohn theorem predicts that, electric dipole transition energy is insensitive to electron–electron interaction [1, 2]. Recently, several researchers have been concerned with vertically coupled dot systems, in which the generalized Kohn theorem is invalid, because of an interdot interaction. The observation of electron–electron interaction can be expected in FIR absorption response of the vertically coupled dot. Partoens *et al.* [3] have treated the electron system in a vertically coupled dot as a magneto-plasma and calculated the coupling effect on a magnetic dispersion of the absorption peaks in the FIR response. The coupling between dots depends on the ratio between the distance and the radius of the dot, if lateral confine potentials are different for both dots. As far as we know, few experimental works have been done with vertically coupled dots. Heitmann [4] observed the breaking of the Kohn theorem. They fabricated a dot array on double quantum wells. The distance between dots is fixed in the sample. It is difficult to change the coupling between dots. We tried to fabricate a vertically aligned double-dot system on double heterostructure samples. In the samples, distance between the two dots depends
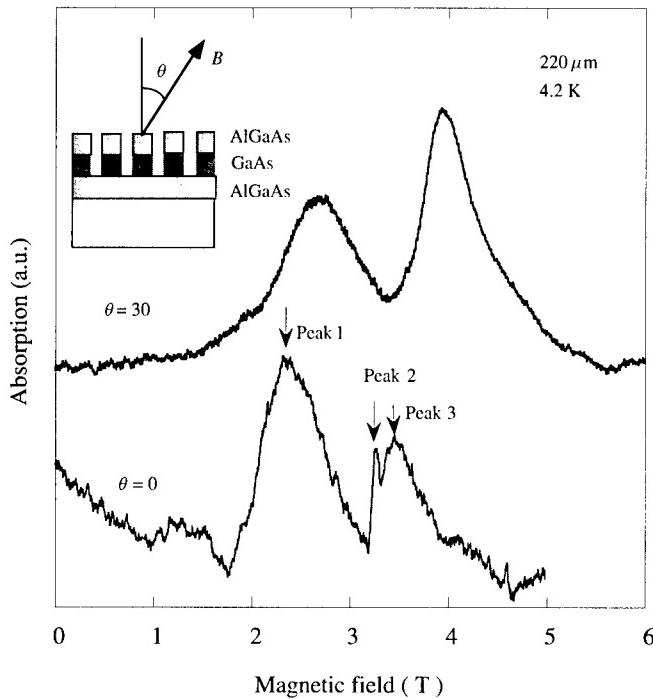
Fig. 1. Far-infrared absorption of a vertically aligned dot array. Two broad peaks (peaks 1 and 3) shift toward higher magnetic field with the angle of the magnetic field. The inset shows the schematic cross-section of a sample.

on the vertical potential profile forming a two-dimensional electron gas (2DEG) at two heterointerfaces. It is expected that the distance between dots can be changed externally. The aim of this paper is to discuss the possibility of formation of the vertically coupled dot system using the double heterostructure samples and potential control by external illumination of light to change the coupling between the dots. In the double heterostructure samples, electronic states split into a symmetric state and an asymmetric state. In the latter state, the wavefunction has peaks in the vicinity of the two heterointerfaces. Ordinary fabrication technique makes vertically aligned double-dot system on the samples. The controlling of the coupling between these two dots modifies the magnetic dispersion in FIR response. We will demonstrate the potential modulation of the double-dot system by pulsed photoexcitation which is useful for systematic investigations of electron–electron interaction in a double-dot system.

2. Experimental set-up

Dot arrays were fabricated on a GaAs/AlGaAs heterostructure using electron beam lithography and wet etching technique. The heterostructure has a 30 nm thickness GaAs layer which is sandwiched between two δ-doped AlGaAs layers. The carrier density and mobility are 5.0×10^{15} m^{-2} and 1×10^5 cm^2 V^{-1} s^{-1}, respectively. The lithographically defined dot size is 2.0 μm. The constructed area of the dot array is about 3×2 mm^2, including some 10^6 dots. For the FIR magneto-optical measurement, the laser lines between 119 and 433 μm were employed. For the confirmation of the two-dimensionality, absorption measurements were done under the tilted magnetic field. Experiments were performed under Faraday configuration at 4.2 K. For pulsed photoexcitation, xenon flash lamp was used. To eliminate the contributions from the substrate, semi-insulating GaAs, the samples were illuminated by infrared light for more than 10 min before the mea-
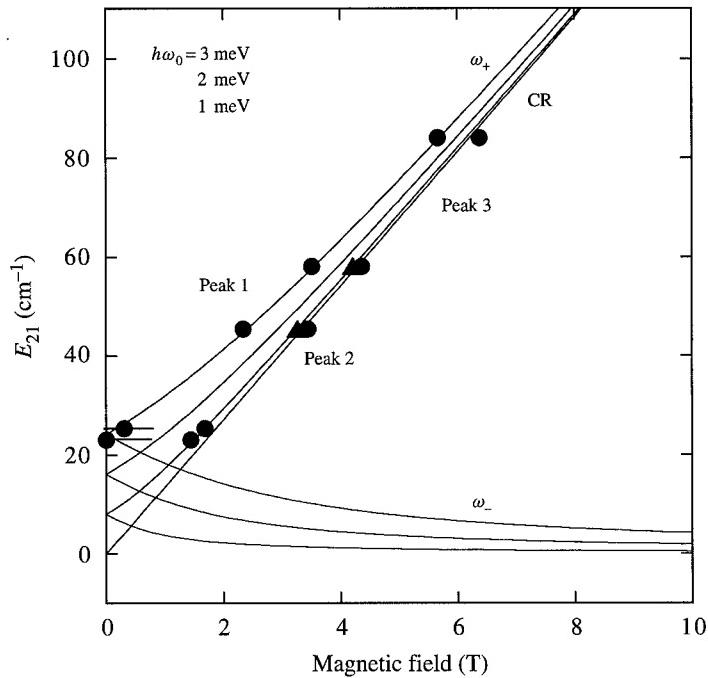
Fig. 2. Magnetic field dependence of the peak position and calculated magnetic dispersion curves for several values of ω_0. Experimental data can be well fitted with the curves for 1 and 3 meV of ω_0.
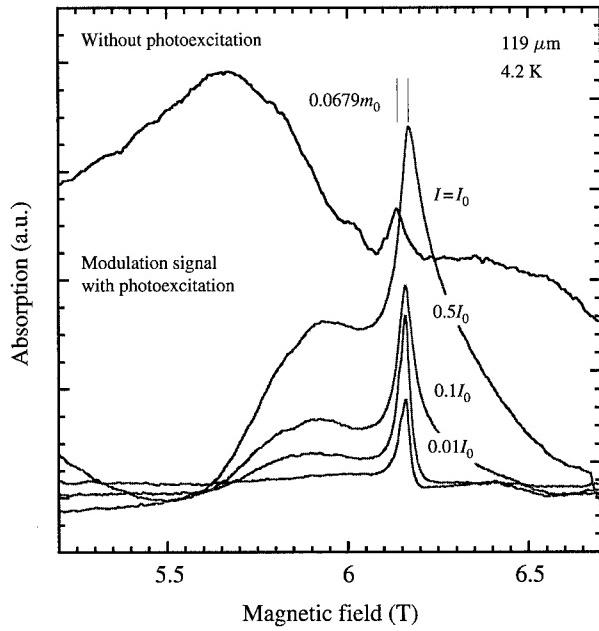
Fig. 3. Modulated absorption signal under bandgap excitation. At the position of peak 1, the characteristics change due to peak shift being recognized.

surements. The illumination also eliminates the persistent photoconductivity effect in the time-resolved measurements.

3. Experimental results

Figure 1 shows a magnetic field dependence of FIR absorption. The measurement was carried out at 4.2 K with 220 μm laser line. Two broad peaks (peak 1 and peak 3) and a sharp peak (peak 2) are observed at 2.25, 3.34 and 3.15 T, respectively. In the tilted magnetic field, the peak positions of the broad peaks shift toward higher magnetic field. The cyclotron radius of electrons at 3 T is 15 nm. Then electrons related to the broad peaks are confined in the narrow region less than 15 nm. This means the broad peaks are due to the 2DEG localized in the heterointerfaces. The difference in the resonance fields of these two peaks are caused by the different lateral confinement of 2DEG. While the peak position of the peak 2 is independent of the direction of the applied magnetic field. Under photoexcitation the peak 2 appears in almost the same magnetic field. For the peak 2, the electron wavefunction extends in the whole GaAs layer. We concluded that the broad peaks are caused by asymmetry state and the sharp peak is related to the symmetry state. Figure 2 shows the magnetic dispersion of transition energy deduced from the peak position in the FIR absorption measurements. The magnetic dispersion depends on a characteristic frequency of the dot ω_0 . Solid lines shows the calculated dispersion curves for 1, 2, and 3 meV of ω_0, respectively. The peak positions for peak 1 and 3 are well fitted by the calculated lines. The characteristic frequencies ω_0 are found to be 3 and 1 meV for peak 1 and peak 3, respectively. Figure 3 shows the modulated absorption signals under photoexcitation. The intensity of peak 2 increases. The peak shift of peak 1 to the higher magnetic field is also observed. Under photoexcitation, impurity cyclotron resonance (ECR) peak also appeared. From the temperature dependence of the FIR absorption signals, it is not a heating effect. This is also evidence of the change in the confined potential by photoexcitation.

4. Discussion

We already observed the peak shift in a single-dot array under photoexcitation [5]. This shift is interpreted as the reduction of the vertical confinement potential. Photoexcited electrons neutralize ionized donors in AlGaAs layer. Space charge generated near the heterointerface decreases. Then photoexcitation may change the depletion layer length as well as surface potential. The peak shift of peak 1 shown in Fig. 3 is caused by the reduction of the vertical confined potential. This reduction changes the extent of the wavefunction of the dot. Then the distance between the dots reduces. It is expected that the coupling between the dots is changed by photoexcitation. We calculated the potential and squared wavefunctions of double heterostructures by a self-consistent finite element method after the manner of Inoue *et al.* [6] The calculation was carried out for various surface potentials and depletion lengths to investigate the effect of the photoexcitation. It is found that the confined potential is reduced and Fermi energy decreases by photoexcitation. The decrease of the Fermi energy due to the photoexcitation may enhance the available final-state number. Thus, the intensity of peak 2 increases. In the vertically coupled dot, coupling between two dots is controlled by the distance between the dots. We observed vertical confined potential modulation by photoexcitation. It is possible that photoexcitation changes the coupling between dots.

References

[1] T. Demel, D. Heitmann, P. Grambow, and K. Ploog, *Localization and Confinement of Electrons*, edited by F. Kuchar, H. Heinrich, and G. Bauer (Springer-Verlag, Berlin, 1990) p. 52.

[2] L. Brey, N. F. Johnson, and B. I. Halperin, Phys. Rev. **B40**, 10647 (1989).

[3] B. Partoens, A. Matulis, and F. M. Peeters, Phys. Rev. **B57**, 13039 (1998).

[4] D. Heitmann, K. Bollweg, V. Gudmundsson, T. Kurth, and S. P. Riege, Physica **E1**, 204 (1997).

[5] K. Fujii, T. Yoshizawa, T. Ohyama, K. Oto, S. Takaoka, K. Murase, and K. Gamo, Jpn. J. Appl. Phys. **38**, 542 (1999).

[6] K. Inoue, H. Sakaki, J. Yoshino, and T. Hotta, J. Appl.Phys. **58**, 4277 (1985).

Evaluation of the effect of fabrication tolerances on the ground-state energy of electrostatically defined quantum dots

M. MACUCCI, G. IANNACCONE

Dipartimento di Ingegneria dell'Informazione, Università degli studi di Pisa, Via Diotisalvi 2, I-56126 Pisa, Italy

C. VIEU, H. LAUNOIS, Y. JIN

L2M, CNRS, Avenue Henri Ravéra 196, F-92225, Bagneux, France

We have investigated the precision achievable, with state-of-the-art lithography, in the fabrication of quantum dots for the realization of quantum cellular automaton cells, and we have compared it with the requirements for proper device operation. Our conclusion is that a simple 'hole-array' approach is not feasible, and that individual tuning of each dot is necessary.

© 2000 Academic Press

Key words: quantum cellular automata, fabrication tolerances.

1. Introduction

One of the most important obstacles to the realization of large-scale circuits based on mesoscopic devices is represented by their sensitivity to fabrication tolerances and to the presence of stray charges. Precision requirements, in particular, are extremely severe for some proposed devices [1], such as those based on the quantum cellular automaton (QCA) concept [2].

We present an experimental and theoretical investigation of the precision achievable in the fabrication of 4-dot cells defined by means of openings in a metal gate evaporated on top of a GaAs/AlGaAs heterostructure. Specifically, our aim has been that of determining the dispersion of the ground-state energy of single electrons in each dot, due to tolerances in the boundaries of the gate openings. Such an investigation has been performed both for the case of single dots and for that of groups of four dots located at the vertices of a square (as in the case of a QCA cell). Gates with arrays of square openings have been fabricated by means of state-of-the-art high-resolution electron-beam lithography and scanning electron microscope (SEM) photographs of the resulting structures have been processed, extracting the contour of each opening. The confinement potential generated by such contours at a depth corresponding to that of the two-dimensional electron gas (2DEG) has then been computed, assuming Fermi level pinning at the semiconductor–air interface, and statistics about the ground-state energy for each dot have been derived and discussed.
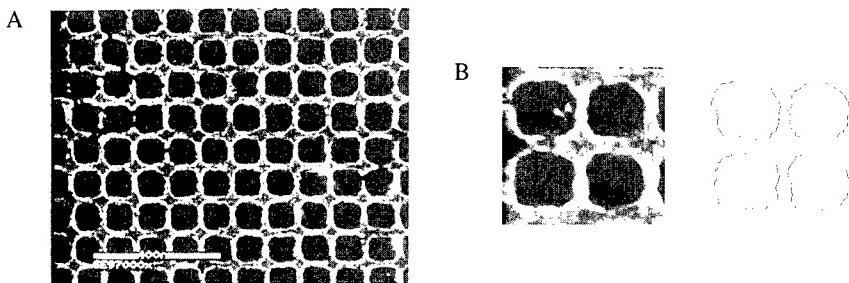
Fig. 1. A, Dot array with 60 holes, with an average hole area of 5961 nm² and a standard deviation of 349 nm². B, Comparison between the SEM image of a cell and the manually extracted contours.

2. Sample fabrication

Samples have been fabricated at L2M on a GaAs/AlGaAs heterostructure consisting of a 20 nm AlGaAs spacer layer, a delta doping layer of Si, of about 6×10^{12} cm^{-2}, an undoped 10 nm AlGaAs layer, and, on top, an undoped 5 nm GaAs cap layer. The 2DEG is therefore at a depth of 35 nm from the surface.

Since our aim was that of evaluating the minimum size fluctuations achievable with state-of-the-art technology, a very high-resolution electron-beam lithography system has been used, based on a modified Philips CM20FEG STEM (scanning transmission electron microscope). A probe size of less than 1 nm, with a current of 12 pA and an electron energy of 200 keV has been selected, and the actual linewidth on the resist has been varied between 10 nm and 25 nm by tuning the dose between 3 nC cm^{-2} and 6 nC cm^{-2}.

The sample has been first spun to deposit a 100 nm layer of PMMA resist and baked at 170 °C before exposure. After exposure, the resist has been developed in MIBK-isopropyl alcohol. A 10 nm titanium and a 20 nm gold layers have then been evaporated on the sample, and, finally, a lift-off process with trichloroethylene solvent has been performed. The main technological difficulty is that of avoiding a poor lift-off.

3. Numerical results

Let us first consider the dot array shown in Fig. 1: it contains 60 holes, whose area has an average of 5961 nm² and a standard deviation of 349 nm².

The contour of each dot has been extracted visually, by manually following the edge of the holes in the digitized SEM image shown in Fig. 1, using the 'xfig' public-domain graphic software. The xfig output file has then been processed with a purposely developed automatic procedure, in order to compute the area of each hole and, for each cell, the confinement potential acting on a 2DEG at a depth of 35 nm. The confinement potential has been obtained implementing the quasi-analytical procedure proposed by Davies and Larkin [3], which allows the determination of the potential produced by an arbitrary polygonal gate on top of a heterostructure, within the approximation of Fermi level pinning at the exposed GaAs surface. The slight variation of permittivity in the heterostructure has been neglected. The holes have been divided into groups of four (corresponding to the four openings needed to define a QCA cell) and the confining potential for each dot has been computed as a result both of all the holes defining the cell and of just the corresponding hole (as if the hole defining it were isolated from the others). Then, the single-particle Schrödinger equation has been solved for each computed confining potential, in order to obtain the ground-state energies.

In Fig. 2, we show the scatter plot of the dot ground-state energy versus the reciprocal of the area, for a dot in a cell (open circles) and for single isolated dots (full circles). In both cases a very good linear fit is obtained: the correlation coefficient is 0.998 for isolated dots, and 0.927 for dots in a cell. In the latter case the ground state energy is lower and more randomized, because of the effect of the holes defining the
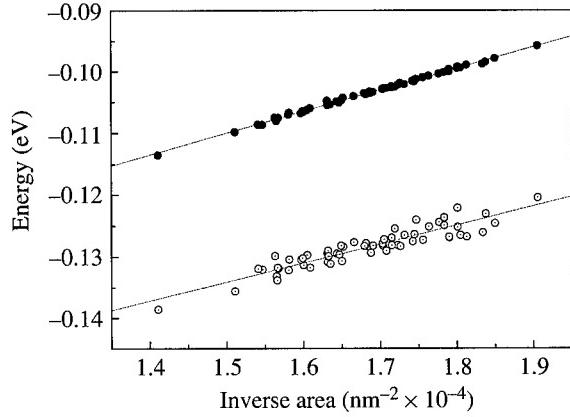
Fig. 2. Ground-state energy versus inverse area for single dots (full circles) and for dots in a cell (open circles), relative to the sample shown in Fig. 1.
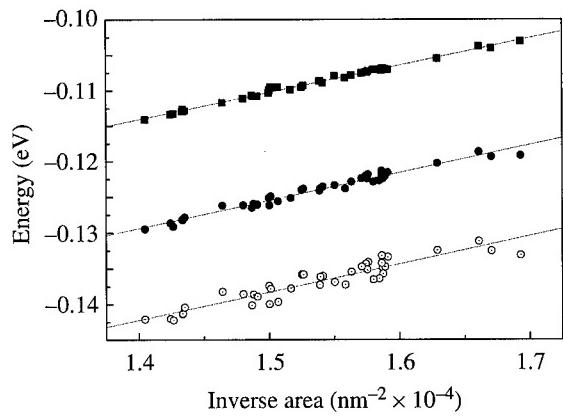
Fig. 3. Ground-state energy versus inverse area of for the dots in a cell with a pitch of 100 nm (open circles), for the dots in a cell with a pitch of 200 nm (full circles), and for single dots (full squares), relative to the second hole-array sample.

rest of the cell, whose area fluctuates independently of that of the hole relative to the dot being considered. The second set of samples is a similar dot array of 40 holes, whose area has an average of 6520 nm^2 and a standard deviation of 295 nm^2.

In Fig. 3 the ground-state energy is plotted versus the inverse area for the second sample. Results are reported for a cell with an interdot distance (measured between the centers of adjacent dots) of 100 nm (open circles), for an enlarged cell obtained spreading the dots apart so that the interdot distance becomes 200 nm (full circles), and for isolated dots (full squares). Also, in this case the average energy decreases and randomization increases as the interaction between adjacent holes is incremented.

As far as the statistics of ground-state energies are concerned, the standard deviation is about 3% of the average value, i.e. about 4 meV, for both sets of data. The operation of a QCA cell will be disrupted if the fluctuation of the ground-state energy in one of the dots is comparable to or larger than the energy splitting between the two configurations [1]. For cells with an interdot separation of 100 nm, obtained in a 2DEG at a depth of 35 nm, and considering a separation between cell centers of 150 nm, we obtain a splitting of only

0.022 meV (the situation would somewhat improve for a heterostructure with a deeper 2DEG: 0.065 meV for a depth of 100 nm). In order to obtain a reasonably reliable QCA array, the standard deviation of the ground-state energy should be no more than one-tenth of the splitting, i.e. 0.0022 meV. It is therefore apparent that the achievable precision is three orders of magnitude less than what would be required, and that, even allowing for technological improvements, a 'hole-array' approach cannot realistically lead to a working QCA circuit. Furthermore, even larger fluctuations of the confinement potential may result from the random distribution of dopants and impurities, particularly in the case of heterostructures in which the delta doping layer is close to the 2DEG. At this stage a tunable design, such as that with several split gates proposed in [1], is better suited to demonstrate the principle of QCA operation.

References

[1] M. Governale, M. Macucci, G. Iannaccone, C. Ungarelli, and J. Martorell, J. Appl. Phys. **85**, 2962 (1999).
[2] C. S. Lent, P. D. Tougaw, and W. Porod, Appl. Phys. Lett. **62**, 714 (1993).
[3] J. H. Davies, I. A. Larkin, and E. V. Sukhorukov, J. Appl. Phys. **77**, 4504 (1995).

Linear and nonlinear transports of coupled quantum dots

SHUJI TSUBAKI, NAOKI FUJITA, FUJIO WAKAYA, YOSHIHIKO YUBA, KENJI GAMO

Department of Physical Science, Graduate School of Engineering Science, Osaka University and Research Center for Materials Science at Extreme Conditions, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan

Series of double quantum dots each with a size around 400×400 nm^2 have been realized by delineating a 2DEG in modulation-doped AlGaAs/GaAs with 100 nm wide Schottky split gates fabricated by an electron-beam lithography and a lift-off technique. The split gate in the middle of the double dot allows us to control interdot coupling widely. The charging diagram obtained from linear transports in the Coulomb blockade regime shows that the isolated dots merge into a single composite dot with increase of interdot coupling. A clear Coulomb staircase has been observed in the double-dot system at a limited high-bias condition.

© 2000 Academic Press

**Key words:** quantum dots, linear transport, nonlinear transport.

1. Introduction

A coupled quantum dot system [1] in the Coulomb blockade (CB) regime has been extensively investigated both theoretically [2–4] and experimentally [5–7]. Depending on the coupling, the coupled dot system shows various mesoscopic and single-electron transports such as stochastic Coulomb blockade [8], peak splitting and charge fluctuations [2–6]. A coherent mode or molecular behavior in the system has been resolved by controlling an interdot coupling [9, 10]. There has been a growing interest in the coupled dot because of its potential application for the single electronics and the quantum computing.

To realize various coupled-dot systems with a tunable interdot coupling, we believe a delineation of a two-dimensional electron gas (2DEG) in a modulation-doped heterostructure with split gates is most suitable since dot sizes, dot potentials, tunnel barrier heights and interdot couplings can be controlled readily and precisely. Such surface split gates have been fabricated by utilizing an electron-beam lithography and a lift-off of gate metals. We investigated coupled dot systems with a variable coupling fabricated by this process to reveal linear and nonlinear transports in CB regime and possible issues of dot integrations.

2. Sample fabrications and measurements

The samples used were multiple quantum dots each with a size of 350×350 nm^2 defined and separated by independently bias-tunable gates. They were fabricated from a MBE-grown, modulation-doped AlGaAs/GaAs, which contained a 2DEG with a sheet concentration of 2×10^{11} cm^{-2} at a depth of 60 nm below the surface. Schottky barrier gates with the minimum width and the minimum separation of 100 nm

Drain 1   **PG4**   Drain 2

PG3   PG5

CG2 Dot 2   Dot 3 CG3

PG2   PG6

CG1 Dot 1

PG1

                    PG7

Source 1   Source 2

1 μm
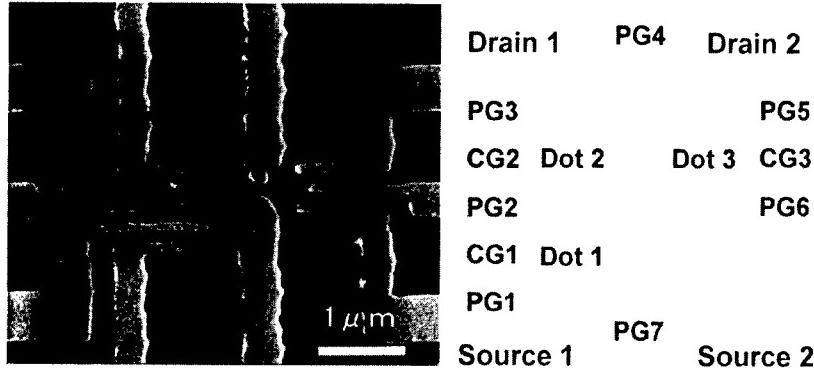
Fig. 1. SEM photograph of the multi-quantum dot sample and a schematic diagram of the configuration of quantum dots and split gates. The series-coupled sample of dot 1 and 2 and source 1 and drain 1 were used in the present experiment with the third dot cut-off by the split gate.

were formed on the top surface of the sample by using a lift-off technique of the deposited metal films of AuPd with a resist mask exposed by electron-beam lithography.

Figure 1 is a scanning electron micrograph of 10 electrostatic gates defining and controlling three quantum dots and a schematic illustration of the configuration of dots and gates. Gates labeled PGi define independently tunable quantum point contacts (QPCs). The outer two QPCs are used to measure the conductance of coupled dot in the CB regime, while the inner QPC controls the interdot coupling. Voltages applied to the side gates CG1 and CG2 are used to change the electrostatic potential of each dot.

In the present experiment, we concentrated on the series-coupled system of dot 1 and dot 2 and tuned the point contact between PG4 and PG7 to cut off the third dot. Transport measurements were carried out at temperatures around 50 mK in a dilution refrigerator by use of a low-noise dc method as well as a conventional lock-in technique.

3. Results and discussion

The double quantum dot is best characterized by measuring its charging diagram. Source–drain currents through the double-dot system were traced by varying the electrostatic potentials of two independent gates CG1 and CG2. To control the interdot coupling, V_{PG2} is manipulated around $-1.20 \sim -1.22$ V and to define the double quantum dot, the following negative voltages were applied: $V_{PG1} = -0.93$ V, $V_{PG3} = -1.3$ V and $V_{PG4} = V_{PG7} = -1.2$ V. Figure 2 represents source currents (I_s) for three different interdot couplings at a source–drain bias of 100 mV in a gray scale, where $I_s < 6$ pA refers to black and $I_s > 10$ pA refers to white. The interdot conductance is a parameter describing the interdot coupling but it is not possible to measure it directly in the double dot in the CB regime. We measured a conductance of the sample with PG2 and PG4 set at the same voltages used for defining the double dot and the other gates grounded. The measured conductances are 0.9, 0.75 and 0.65 G_Q where $G_Q = 2e^2/h$ for V_{PG2} of -1.205, -1.215 and -1.22 V, respectively, and using these values, we define the strong, the intermediate and the weak coupling cases of the present experiment. These conductances may be different from the interdot conductances in the coupled double dot since reverse biased gates near PG2 affect the interdot transport and are expected to be higher compared with those in the coupled double dot.

In Fig. 2A, a roughly rectangle but somewhat skewed array of current peaks was observed. This suggests that each dot is well defined and the coupling is weak because the current can flow through a double dot only at certain settings of the gate voltages, where the CB is lifted for both dots simultaneously. Each tetragon
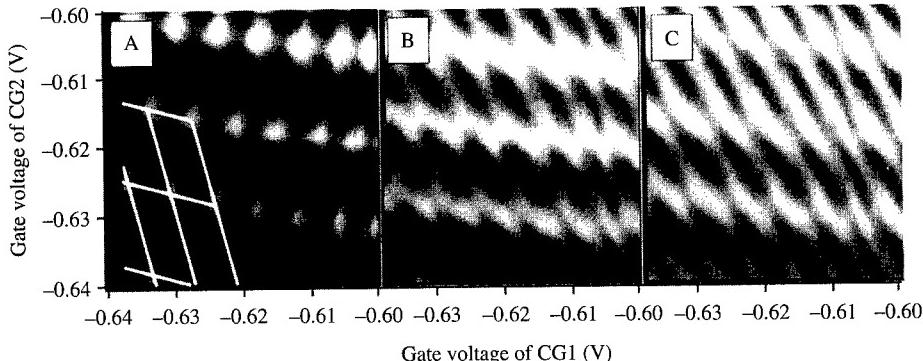
Fig. 2. Gray-scale plots of source current I_s at drain voltage $V_d = 100$ mV as a function of V_{CG1} and V_{CG2} for three different interdot couplings controlled by V_{PG2}; A, $V_{PG2} = -1.22$ V, B, $V_{PG2} = -1.215$ V and C, $V_{PG2} = -1.205$ V. On the interdot conductance, see the text. Some tetragons shown by solid white lines as a guide for the eye in A mark regions where a particular configuration of excess electrons in the dot system is stable and there is no charge transport due to the CB.

defined by current peaks marks a region where a particular configuration of excess electrons in each dot of the system is stable and there is no charge transport due to the CB. Some of them are shown in Fig. 2A by solid white lines as a guide for the eye. The gate capacitance of each dot is estimated from this figure as $C_{CG1} = 27.1$ aF and $C_{CG2} = 12$ aF, which agreed well with those determined from the Coulomb oscillation of the single dot. The interdot capacitance is expected to split the current maximum into two peaks [5, 7] but it is not clear, probably because an interdot capacitance is rather small in the present case. In addition, current peaks are blurred because of a finite source drain bias and thermal smearing.

With increasing coupling, a drastic change occurred in the charging diagram from the point array in Fig. 2A. In Fig. 2B, current maxima split and extend to form a hexagon array as observed previously [5, 6]. Due to increased the coupling, CB was partially lifted and appreciable currents were observed along the hexagonal cell, which defines the region of total charge quantization. In Fig. 2C, the gray-scale plot shows an array of nearly parallel lines. This suggests that the double-dot system merged into a single large dot due to the strong coupling. In this sample, only the total charge in the system is quantized in the region between the parallel lines in Fig. 2C and the quantization of charges in each dot is destroyed. These results suggest that it is possible to control widely an interdot coupling in the multiple quantum dot system with split gate.

Influences of interdot coupling on Coulomb gaps and Coulomb oscillations were investigated. A typical example for the intermediate coupling is shown in Fig. 3, which shows the source current I_s versus drain voltage V_d traces as a function of gate voltage and the gray-scale representation of I_s on V_d–V_g plane. In these measurements, two gates were tied and swept with the same voltage. We observed a series of Coulomb diamonds with two different sizes and it seems that large and small diamonds appeared alternatively. This corresponds to the peak splitting in the Coulomb oscillation that was expected and observed in the double dot with a finite interdot tunnel coupling [2, 3, 5, 6]. Therefore, the charge quantization in each dot is destroyed and an excess charge is shared between dots in the intermediate coupling sample. It is different from the characteristics of the single dot or the weakly coupled dot.

Nonlinear characteristics in double-dot system has been reported to depend on the interdot coupling and the Coulomb staircase is expected to become evident with coupling [4, 6]. Figure 4 shows I_s versus V_d characteristics as a function of gate voltage for the sample at a further weak coupling compared with those in Fig. 2 and the gray-scale representation of I_s on V_d–V_g plane. We observed some plateaus with a step width around 0.5 mV in the I_s–V_d characteristics outside the CB regions, which were probably the Coulomb staircase because of their regularity and step-width. These structures seemed to be regular for varying drain bias for small
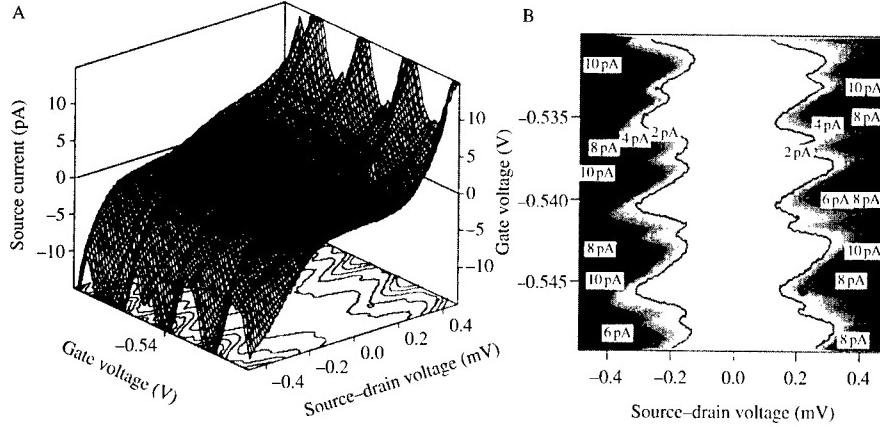
Fig. 3. A, Source current I_s versus drain voltage V_d characteristics of double dots as a function of gate voltage V_g for the intermediate interdot coupling (Fig. 2B) and B, the gray-scale representation of I_s on the V_d–V_g plane.
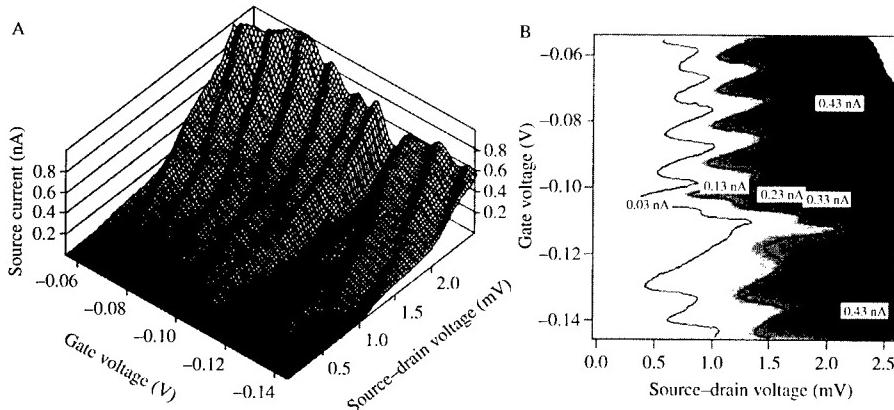
Fig. 4. A, Source current I_s versus drain voltage V_d characteristics of a double dot with a weak coupling as a function of gate voltage at higher drain bias and B, the gray-scale representation of I_s on the V_d–V_g plane.

Coulomb diamonds but disappeared for a large Coulomb diamond, which might suggest that asymmetry of the tunnel barrier is essential in being able to observe the Coulomb staircase in I_s–V_d characteristics.

4. Summary

We fabricated coupled quantum dots with tunable interdot coupling on AlGaAs/GaAs modulation-doped heterostructure by using a split-gate technique and investigated their linear and nonlinear transports in the CB regime. The charging diagrams for the coupled dots with different coupling strengths were obtained, which showed a drastic change from nearly perfectly isolated dots to a composite single dot with increased coupling. A clear Coulomb staircase was observed in the limited coupled dot.

References

[1] L. P. Kouwenhoven *et al.*, *Mesoscopic Electron Transport*, NATO ASI Series, edited by L. L. Sohn, L. P. Kouwenhoven, and G. Schon (Kluwer, Dordrecht, 1997) Vol. E345.

[2] K. A. Matveev, L. I. Glazman, and H. U. Baranger, Phys. Rev. **B53**, 1034 (1996); **B54**, 5637 (1996).

[3] J. M. Golden and B. I. Halperin, Phys. Rev. **B53**, 3893 (1996).

[4] R. Kotlyar and S. Das Sarma, Phys. Rev. **B56**, 13235 (1997).

[5] F. R. Waugh, M. J. Berry, C. H. Crouch, C. Livermore, D. J. Mar, R. M. Westervelt, K. L. Campman, and A. C. Gossard, Phys. Rev. **B53**, 1413 (1996); C. Livermore *et al.*, Science **274**, 1332 (1996).

[6] C. H. Crouch, C. Livermore, F. R. Waugh, R. M. Westervelt, K. L. Campman, and A. C. Gossard, Surf. Sci. **361/362**, 631 (1996); C. H. Crouch, C. Livermore, R. M. Westervelt, K. L. Campman, and A. C. Gossard, Appl. Phys. Lett. **71**, 817 (1997).

[7] D. Dixon, L. P. Kouwenhoven, P. L. McEuen, Y. Nagamune, J. Motohisa, and H. Sakaki, Phys. Rev. **B53**, 12625 (1996); Surf. Sci. **361/362**, 636 (1996).

[8] I. M. Ruzin, V. Chandrasekhar, E. I. Levin, and L. I. Glazman, Phys. Rev. **B45**, 13469 (1992).

[9] R. H. Blick, R. J. Haug, J. Weis, D. Pfannkuche, K. v. Klitzing, and K. Eberl, Phys. Rev. **B53**, 7899 (1996); R. H. Blick, D. Pfannkuche, R. J. Haug, K. v. Klitzing, and K. Eberl, Phys. Rev. Lett. **80**, 4032 (1998).

[10] T. Fujisawa, T. H. Oosterkamp, W. G. van der Wiel, B. W. Broer, R. Aguado, S. Tarucha, and L. P. Kouwenhoven, Science **282**, 932 (1998).

Tuning of surface boundary conditions for the 3D simulation of gated heterostructures

G. IANNACCONE, M. MACUCCI, E. AMIRANTE

Dipartimento di Ingegneria dell'Informazione, Università degli studi di Pisa, Via Diotisalvi 2, I-56126 Pisa, Italy

Y. JIN, H. LANOIS, C. VIEU

L2M, CNRS, Avenue Henri Ravéra 196, F-92225, Bagneux, France

(Received 9 February 2000)

We propose a method for treating the boundary conditions at the exposed surface of semiconductor nanostructures, and compare the results from simulations based on such a method with experimental measurements on test devices defined electrostatically by metal gates on AlGaAs/GaAs heterostructures. In particular, we show that the pinch-off voltage of quantum point contacts realized with split gates can be reasonably reproduced, provided the lithographic gap is small enough.

© 2000 Academic Press

Key words: surface states, heterostructures, split gates.

1. Introduction

In the simulation of semiconductor nanostructures the issue of boundary conditions at the exposed surface has often been neglected or treated with drastic approximations that provide only quantitative agreement with the experimental data. Notable exceptions are, for example, the work by Larkin and Davies [1], where the results from Dirichlet and Neumann boundary conditions are discussed, and that by Chen and Porod [2], who considered a distribution of surface states and computed their occupancy solving the Poisson equation in the air above the semiconductor.

We have developed a 3D multi-grid nonlinear Poisson solver for the determination of the confinement potential within a GaAs/AlGaAs nanostructure due to the action of metal gates at the surface, for an arbitrary layer structure, and we have fabricated full and split gates of several different sizes.

Depletion of the two-dimensional electron gas (2DEG) at the heterointerface is strongly dependent on the choice of boundary conditions at the exposed surface and the main focus of our work has been on devising a proper treatment for such a surface. With Dirichlet boundary conditions (i.e. assuming perfect Fermi level pinning) the voltage that needs to be applied to the gates to completely pinch off a quantum point contact is usually larger than what is experimentally measured, due to the screening action of the surface. On the other hand, Neumann boundary conditions with zero electric field at the interface [3] lead to results that are not in good agreement with the measurements, either.
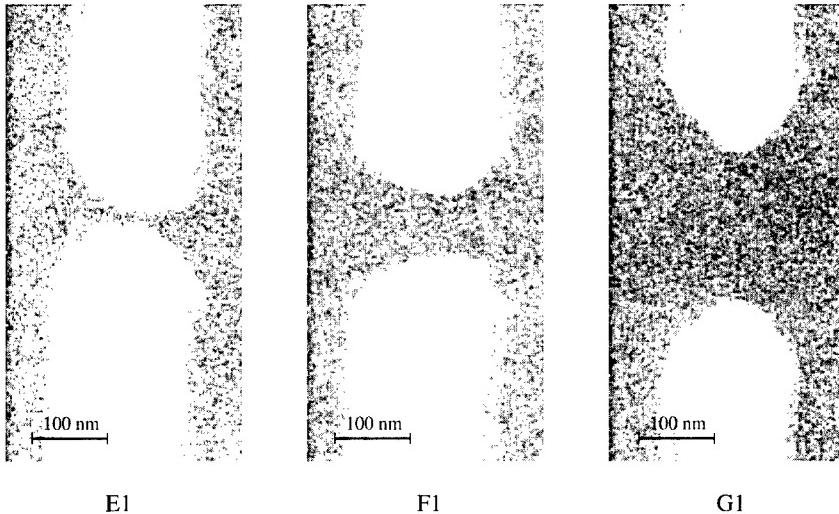
Fig. 1. SEM photographs of the split gate structures E1, F1, and G1.

2. Approach

Our approach consists of performing an initial calculation with all gates grounded and assuming Fermi level pinning at the exposed surface. The difference E_{pin} between the conduction band edge E_C and the energy at which the Fermi level is actually pinned is determined by fitting the experimental sheet charge density in the 2DEG at equilibrium. From the Poisson solver we determine the electric field at the semiconductor–air interface and therefore the surface charge density.

We then proceed considering that at very low temperatures (≈ 4.2 K) the surface charge remains substantially unchanged when nonzero voltages are applied to the gates for normal device operation. In addition, since at the surface there is an abrupt change of the dielectric constant from GaAs ($\epsilon_r = 12.9$) to air, we can reasonably translate the assumption of constant surface charge into that of constant electric field at the GaAs surface. Therefore, the previously determined value of the electric field is used as a Neumann boundary condition for a new solution of the nonlinear Poisson equation with the proper gate voltages. In such a way, the domain of the 3D Poisson equation is limited to the semiconductor, and does not have to be extended to the air above the surface.

3. Experiment

Samples have been fabricated on an AlGaAs/GaAs heterostructure with the following layer sequence: undoped GaAs substrate, 20 nm thick undoped AlGaAs, silicon delta doping with a dose of 6×10^{12} cm^{-2}, 10 nm thick undoped AlGaAs, 5 nm GaAs cap layer. A two-dimensional electron gas is obtained at a depth of 35 nm, with a measured sheet electron density $n_s = 4.9 \times 10^{11}$ cm^{-2} at the temperature of 4.2 K.

Several gate layouts have been patterned by electron beam lithography and successive lift-off. In particular, single gates with a width of 80, 160, 220, and 540 nm, and split gates with a width of 120 nm and different lithographic gaps have been realized: 60, 120, 200 nm for the samples E1, F1, and G1, respectively, which are shown in Fig. 1.

It is important to point out that in order to improve gate adhesion on the surface, a slight etching step (≈ 3–4 nm) precedes metal evaporation, so that the gates are actually deeper than the exposed surface. This has a significant effect on the pinch-off voltage of the full gates, and a much less relevant effect on the pinch-off
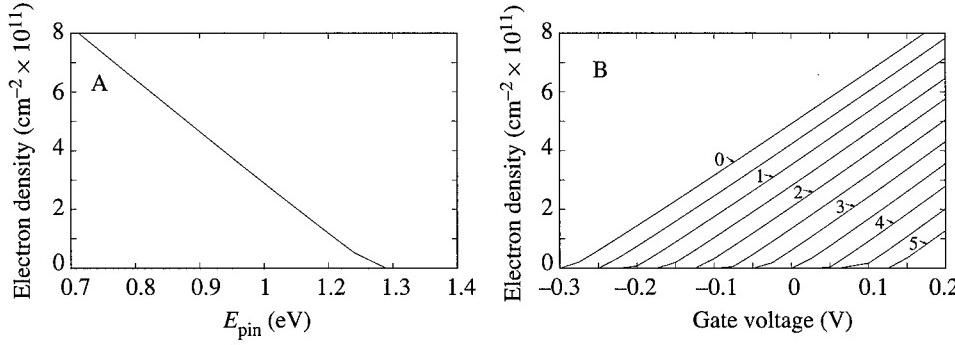
Fig. 2. A, Electron density in the 2DEG as a function of E_{pin}; B, electron density as a function of the gate voltage for different etching depths (nm) with $E_{pin} = 0.74$.
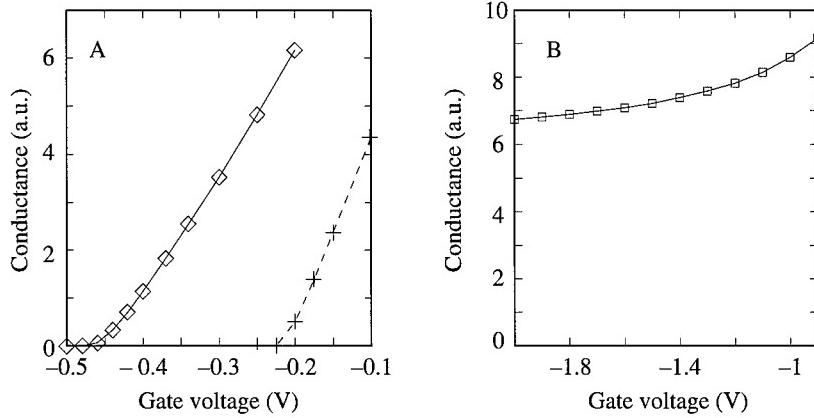
Fig. 3. Theoretical conductance of constrictions E1, F1 (left) and G1 (right) as a function of the gate voltage.

voltage of the split gates. In this paper, we focus on the split gate structures, and, among the full gates, we consider only the one with the largest width, whose pinch-off value should be essentially equal to that due to an infinite metal plane at the surface of the heterostructure.

4. Results and discussion

With a nonlinear 1D Poisson solver we have first computed the value of n_s at equilibrium as a function of E_{pin} (Fig. 2A), in order to obtain the value of E_{pin} corresponding to the measured n_s ($E_{pin} = 0.855$ eV). Such a value is well within the range of values for E_{pin} found in the literature [4]. The corresponding electric field at the surface is $\mathcal{E} = 0.772$ MV cm^{-1}. We have then computed the sheet electron density in the 2DEG as a function of the gate voltage for different values of the etching depth (Fig. 2B): an etching depth of 1 nm corresponds to a shift of about 70 mV for the pinch-off voltage. The experimental pinch-off values for the 540 nm wide single gate are 60 and 173 mV, which are well fitted by an etching depth of 4 and 6 nm, respectively.

We have then solved the 3D nonlinear Poisson equation for the split gates. The actual shape of the gates has been obtained from the SEM photographs shown in Fig. 1. We have assumed an etching depth of 4 nm under the gates and for each gate voltage we have computed the electron density, the conduction band edge profiles,

and the Drude conductance, assuming a local conductivity $\sigma(\mathbf{r}) = q\mu n(\mathbf{r})$, where q is the electron charge, μ the mobility, n the electron density. In Fig.3 the conductance is plotted as a function of the gate voltage for the three samples. The computed pinch-off voltages can be obtained directly from Fig.3: they are -0.22 and -0.46 V for E1 and F1, respectively, while there is no pinch-off up to -2 V for G1. The experimental values are -0.23, -0.39, -0.97 V for E1, F1, and G1, respectively. The agreement is reasonable for the split gates with the smaller gaps. In the case of sample G1, the proposed approach fails. We can identify two possible reasons for such a behavior: first, the approximation of constant charge and constant electric field may not be valid for larger exposed areas; furthermore the simulation should take into account the complete gate layout, since the capacitive coupling between the constriction and the gates, in the case of wide gaps, could have a substantial contribution from wide gate regions far from the tips.

References

[1] I. A. Larkin and J. H. Davies, Phys. Rev. **B52**, R5535 (1995).
[2] Minhan Chen and W. Porod, J. Appl. Phys. **78**, 1050 (1995).
[3] M. Stopa, Phys. Rev. **B48**, 18340 (1993).
[4] S. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981) pp. 270–279.

Transport in split-gate silicon quantum dots

A. GUNTHER, M. KHOURY, S. MILIČIĆ, D. VASILESKA, T. THORNTON, S. M. GOODNICK

Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-5706, U.S.A.

We report on the transport properties of novel Si quantum dot structures with controllable electron number through both top and side gates. Quantum dots were fabricated by a split-gate technique within a standard MOSFET process. Four-terminal dc electrical measurements were performed at 4.2 K in a liquid helium cryostat. Strong oscillations in the conductance through the dot are observed as a function of both the top gate bias and of the plunger bias. An overall monotonic and quasi-periodic movement of the peak conductance is observed which is believed to be associated with the bare level structure of the electronic states in the dot coupled with the Coulomb charging energy. Crossing behavior is observed as well, suggestive of either many-body effects or symmetry breaking of the dot states by the applied bias.

© 2000 Academic Press

Key words: silicon quantum dots, transport, split-gate.

1. Introduction

There is considerable interest in quantum dots, as they represent the ultimate reduction in the dimensionality of a semiconductor device. In addition, it is hoped that these devices can extend the observations of single-electron tunneling [1] into the semiconductor device realm where they can be coupled to normal transistors. While there have been several observations of single-electron behavior in GaAs heterostructures, efforts in Si-based devices have been limited to either lithographically defined dots [2, 3], or devices which have relied upon accidental definition of dots [4, 5]. Here, we describe the fabrication and measurements of a dual gate-defined quantum dot, which is embedded within a Si metal-oxide-semiconductor (MOS) field-effect transistor. The dot is formed in the inversion layer created by the top gate, with its lateral definition being provided by two side gates embedded within the gate oxide, thus allowing independent control of both the two-dimensional electron gas density (2DEG) adjacent to the dot, as well as the shape of the dot itself.

2. Fabrication

Quantum dots of differing geometry were fabricated by a split-gate technique within a MOSFET structure using the process described by Khoury *et al.* [6] as shown in Fig. 1. Basically, a channel-stop region is formed by implanting a p^- (100) Si wafer with boron. A thin gate oxide is grown over which narrow chromium gates are defined using e-beam lithography. A field oxide and top inversion gate are deposited for independent control of the dot occupancy.
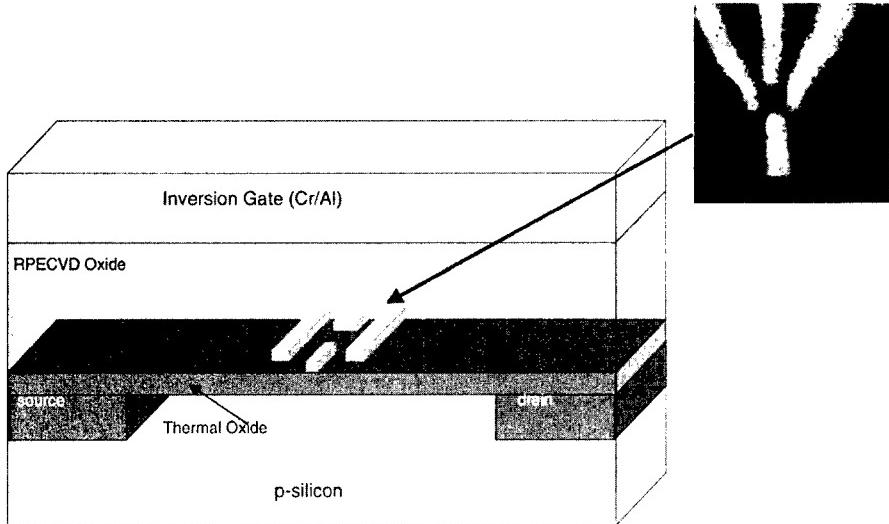
Fig. 1. Schematic diagram of the multilayer Si quantum dot structure. The inset shows the electron micrograph of the depletion gate structure for a 200×200 nm^2 dot.

In the inset of Fig. 1, we show the electron micrograph of a typical device studied here. The dot itself is defined by the side gate pattern, which depletes electrons under the side gates when they are negatively biased, leaving an island of electrons. The input and output quantum point contacts, as well as the plunger gates, are independently biased.

3. Experimental results

Four-terminal dc electrical measurements were performed at 4.2 K in a liquid helium cryostat. Equilibrium measurements were performed by applying a small source–drain excitation. Strong oscillations in the conductance are observed in structures with both normal and with overlapping constrictions. Typical oscillations are shown in the plot in Fig. 2, which are suggestive of Coulomb oscillations due to the filling of the dot one electron at a time. In the figure, the top gate bias was swept for fixed depletion gate. In all measurements, the device is operating within the tunneling regime, where the resistance of the input and output constrictions is much greater than 25 kΩ (the inverse of $G_0 = e^2/h$, the fundamental conductance). Splitting of conductance peaks is observed, which could be due to breaking of the valley degeneracy in Si, as well as the normal spin degeneracy of the states due to the Coulomb charging energy.

Figure 3 shows the position of the conductance peaks observed as a function of both inversion gate and side gate bias for the 200×200 nm^2 symmetric structure shown in the micrograph of Fig. 1. These results correspond to a different device than that of Fig. 2, hence the voltage scales are different. Here all the side gates are tied together to a single voltage source. A monotonic decrease in peak position in terms of depletion (side) gate bias with top gate (inversion) gate bias is evident, which follows from the effect of the top gate in increasing the Fermi energy of the 2DEG surrounding the dot, hence lowering the threshold for conduction in terms of depletion gate bias. In Fig. 3, we have plotted both the major peaks (solid circles) as well as shoulders or minor peaks indicated by open squares.

Overall, the peaks lie on a background of increasing conductance starting with the minimum in the lower left-hand corner, and increasing to a maximum at the upper right-hand side, as indicated by the background shading (darker regions correspond to regions of lower conductance). This trend in conductance reflects
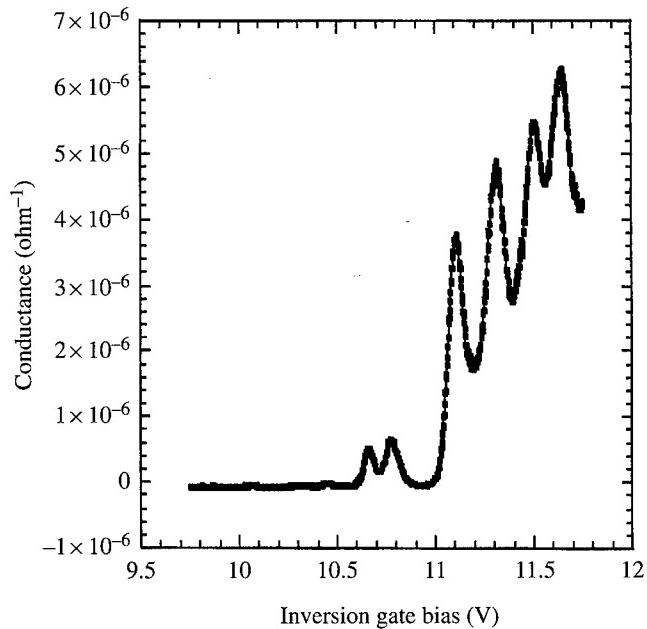
Fig. 2. Single electron oscillations in the quantum dot structure of Fig. 1.
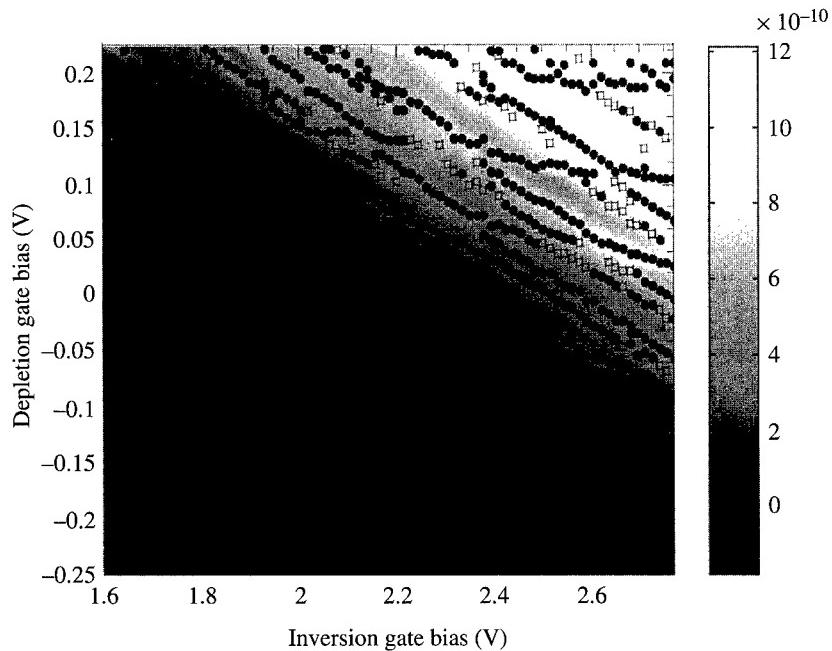

Fig. 3. The peak (solid points) and shoulder (open squares) positions of conductance peaks for various top (inversion) and side (depletion) gate biases for a 200×200 nm^2 dot structure.

the decreased resistance of the quantum point contacts forming the entrance and exit barriers as the 2DEG density is increased, and the side gates open up the dot. In addition to the monotonic behavior observed in the evolution of the peak positions with V_{tg} and V_{sg}, there is strong evidence for crossing behavior, as one 'row' of peaks appears to crossover to an adjacent row at several points in the diagram (for example at $V_{tg} = 2.2$, $V_{sg} = 0.075$ and 0.15 V). Such behavior is even more evident in data in which only one of the gates is swept (for example, the plunger gate).

4. Discussion of results

The crossing behavior shown in Fig. 3 is reminiscent of similar crossover behavior observed in the capacitance spectroscopy of GaAs/AlGaAs quantum dots reported by Zhitenev *et al.* [7]. There, such crossing behavior was attributed to a localization–delocalization transition of the states in the dot as increasing electrons filled it, which is essentially a many-body effect due to the other electrons in the system.

An alternative explanation of the crossing behavior observed in the present system is related to the electronic states of a square symmetry dot, if it is deformed asymmetrically from one side. In a square dot, assuming that the depletion potential results in a quasi-harmonic potential, the allowed states may be approximated by

$$E_{n,m} = \hbar\omega_x(V_{tg}, V_{sg})\left(n + \frac{1}{2}\right) + \hbar\omega_y(V_{tg}, V_{sg})\left(m + \frac{1}{2}\right) + E_0(V_{tg}, V_{sg}),$$

where ω_x and ω_y represent the confinement potentials in the x and y directions, and E_0 is the background potential which changes monotonically with top and plunger bias. Obviously, ω_x and ω_y are also both functions top and side gate bias. If $\omega_x = \omega_y$, degeneracies occur in the energy spectrum as one would expect for a two-dimensional harmonic oscillator, i.e. the levels are N-fold degenerate corresponding to the Nth 'shell' of states. If the side gate bias changes for example ω_y much more than ω_x, different slopes are expected as a function of top and plunger gate bias for the states arising from confinement in different directions, giving rise to a crossing of states associated with confinement in different directions.

The determination of the exact behavior of the dot states with bias requires fully self-consistent analysis of the eigenstates in this three-dimensional structure. Such studies are presently underway, and indicate that such symmetry breaking by the applied bias may account for crossing behavior when asymmetric bias is applied [8], although less so for the present case in which all the gates are tied together, suggestive that multi-particle effects still play an important role in determining the exact spectrum of states in the dot.

References

[1] See e.g., D. K. Ferry and S. M. Goodnick, *Transport in Nanostructures* (Cambridge University Press, Cambridge, U.K., 1997).

[2] L. Guo, E. Leobandung, L. Zhuang, and S. Y. Chou, J. Vac. Sci. Technol. **B15**, 2840 (1997).

[3] L. Zhuang, L. Guo, and S. Y. Chou, Appl. Phys. Lett. **72**, 1205 (1998).

[4] H. Ishikuro and T. Hiramoto, Appl. Phys. Lett. **71**, 3691 (1997).

[5] A. C. Irvine, Z. A. K. Durani, H. Ahmed, and S. Biesemans, Appl. Phys. Lett. **73**, 1113 (1998).

[6] M. Khoury, M. J. Rack, A. Gunther, and D. K. Ferry, Appl. Phys. Lett. **74**, 1576 (1999).

[7] N. B. Zhitenev, M. Brodsky, R. C. Ashoori, L. N. Pfeiffer, and K. W. West, Science **285**, 715 (1999).

[8] S. Miličić, F. Badrieh, D. Vasileska, A. Gunther, and S. M. Goodnick, Superlattices and Microstructures **27**, 377 (2000).

3D modeling of silicon quantum dots

S. N. MILIČIĆ, F. BADRIEH, D. VASILESKA†, A. GUNTHER, S. M. GOODNICK

Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-5706, U.S.A.

We present results of full 3D self-consistent simulations of the energy spectrum in silicon-based symmetric quantum dots. Numerically derived conductance peak dependence upon the depletion and top gate biases closely resembles the experimentally measured ones, suggesting that conductance peak is measured when some discrete energy level in the dot coincides with the Fermi level. Electron wavefunction mode mixing is observed when atomistic description of the impurity distribution in the semiconductor was used.

© 2000 Academic Press

Quantum dots represent the ultimate reduction in the dimensionality of a semiconductor device. While there have been several observations of single-electron behavior in GaAs heterostructures, efforts in Si-based devices have been limited to either lithographically defined dots [1, 2] or using self-assembly techniques [3–5]. To overcome these difficulties, our group at ASU has recently developed a novel technique for the fabrication of Si quantum dot structures in which a controllable number of electrons in the dot is achieved through appropriate biasing of the top and side (depletion) gates. The top gate controls the inversion layer density while the lateral definition of the dot is provided by side gates embedded within the gate oxide (Fig. 1A). It was also observed experimentally that, by varying the top and depletion gates bias, conductance peaks appear that were believed to map the details of the energy level structure within the dot [6].

Herein we present simulation results obtained with our 3D Schrödinger–Poisson solver that show clear correlation between the conductance peaks with the energy level spectrum in the dots. We use the implicitly restarted Arnoldi method for the solution of the 3D Schrödinger equation in the dot region. More precisely, the Lanczos/Arnoldi factorization is used to construct an orthogonal basis for a Krylov subspace that provides a way to implement the projection numerically. Implicit restarting is used to overcome the intractable storage and computational requirements in the original Lanczos/Arnoldi method. As such, this method allows us to calculate only the lowest eigenvalues and the corresponding wavefunctions that describe the occupied states in the dot at low temperature. We use Sturm sequencing, combined with bisection and inverse iteration, to find the eigenvalues and the corresponding eigenfunctions of a tridiagonal symmetric matrix, that arises from the finite-difference discretization of the 1D Schrödinger equation in the leads (quasi two-dimensional electron gas—Q2DEG). Matching of the wavefunctions from the solutions of 1D and 3D Schrödinger equation is done by setting the wavefunction to zero in between. For the solution of the 3D Poisson equation, we use the incomplete LU decomposition (ILU) method. In our theoretical model we also take into account the partial ionization of the impurity atoms, which is crucial for low-temperature device operation.

†E-mail: srdjan@asu.edu, vasilesk@imap2.asu.edu
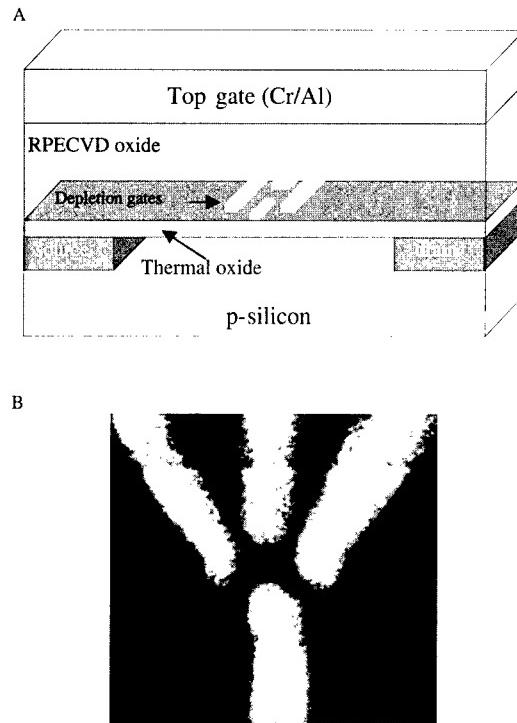
Fig. 1. A, Schematic representation of the Si quantum dot. B, SEM micrograph of the depletion gate pattern.

In Fig. 2A and B we show 25 lowest eigenvalues in the dot region of the device versus applied depletion gate bias when all depletion gates are tied together. The only difference between the results shown in Fig. 2A and B is the applied top gate voltage. From the results shown, it is clear that larger top gate voltage moves all energy levels within the dot down, while applying larger negative depletion gate voltage moves energy eigenvalues up. This behavior is expected and it proves that our simulator is working properly. Note that the energy eigenvalue dependence on the applied depletion gate voltage is approximately linear. We use this linear dependence to extrapolate simulation results for the eigenvalue spectrum into region of applied depletion gate voltages where eigenvalues cross the Fermi level. The Fermi level is taken to be the reference level ($E_F = 0$) in all our simulations.

In Fig. 3A and B we present a top view of the conduction band edge and electron density in the plane parallel to the semiconductor–oxide interface. The confining potential closely resembles fabricated device layout (see Fig. 1B) and it also indicates whether the system (dot) is closed or not at the quantum point contacts (QPC). From the results shown in Fig. 3B it is clear that electrons are completely confined within the dot. The number of electrons in the dot can be varied with both top and depletion gate biases. If we assume that at low temperature only those energy levels below Fermi level are occupied, this number ranges between 15 and 40 electrons for simulated top and depletion gate biases. It should be noted here that with more electrons in the dot, convergence of the program is more stable, with less numerical oscillations, but it takes longer to finish (higher number of electrons inside the dot increases the time required to solve for the energy spectrum in the dot).

The experimentally derived and the simulation results for the conductance peaks of a symmetric 200 × 200 nm quantum dot with all depletion gates tied together are shown in Figs 4A and B, respectively. The procedure used to identify the conductance peaks with the energy-level spectrum in the dot consists of the
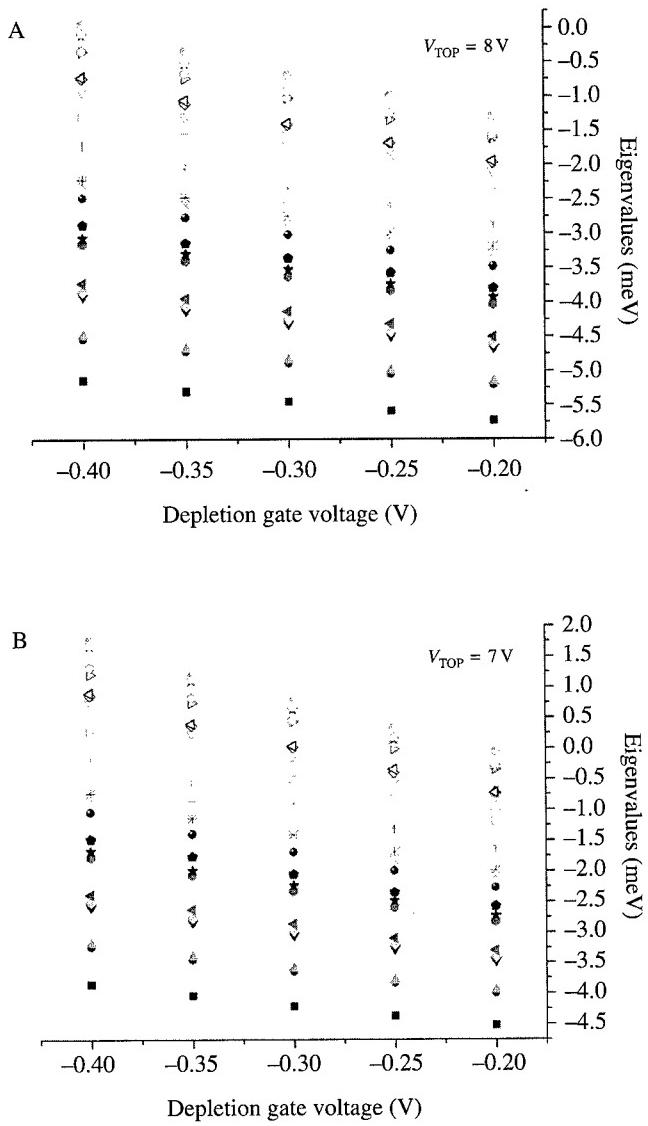
Fig. 2. Energy spectrum versus applied depletion gate voltage. The top gate voltage is fixed at A, 8 V and B, 7 V.

following steps. We first generate a set of solutions that give the variation of the dot energy level spectrum as a function of the depletion bias (with fixed inversion, or top gate bias), like the ones shown in Figs 2A and B. These data sets are then used to derive a parametric expression for the variation of each energy level with inversion and depletion gate biases. We also keep track of the 2DEG sheet-charge density n_s variation with the applied top gate voltage. This essentially tells us the relative position of the ground state of the 2DEG with respect to the Fermi level. A conductance peak is then identified to occur for a combination of depletion-gate and inversion-gate biases for which an energy level in the dot crosses the Fermi level. Comparing the results shown in Fig. 4A and B we observe almost identical slopes for the conductance peaks variation upon the top and depletion gate biases. However, the theoretically predicted gap is not seen experimentally and this
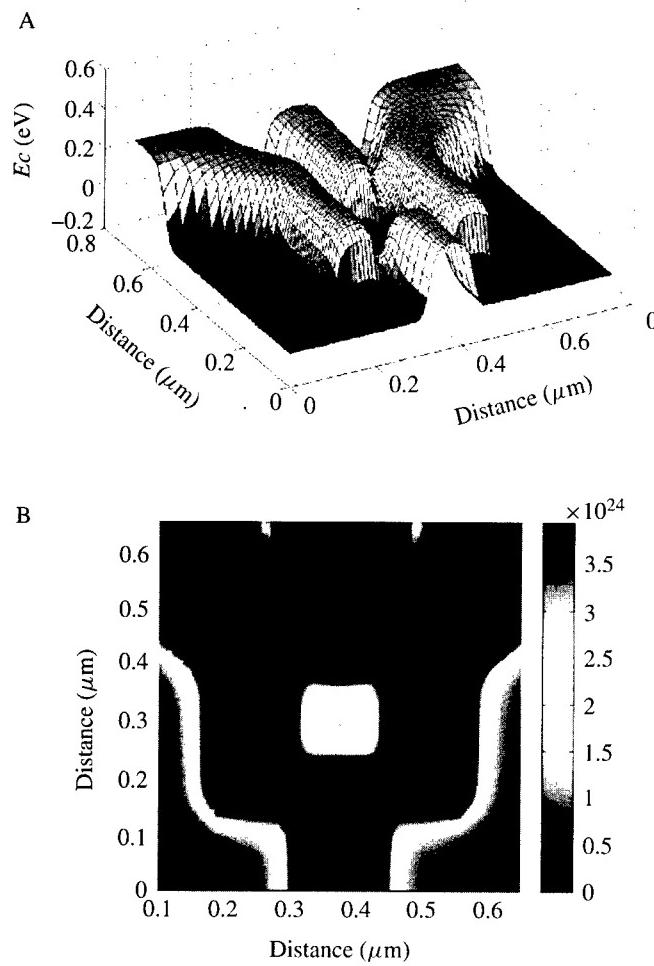
Fig. 3. A, Top view of the conduction band edge. B, Electron density in the plane parallel to the semiconductor–oxide interface.

needs further investigations and refinements of our model to include level interactions that can cause such a gap.

We want to point out that the top gate voltage in both Fig. 4A and B is such that the inversion charge density of the 2DEG varies within the same range in both the experimental and the simulated device. The depletion gate voltage range is also chosen to be the same in both figures. It is important to note that the difference of approximately 5 V for the required inversion gate bias V_G to achieve the desired inversion charge density n_s in the channel, is due to the existence of oxide charges in the fabricated device, which shift the threshold voltage, i.e. the top gate voltage V_G for which the surface is inverted. The oxide charges are not included in our theoretical model, but can be easily corrected for, since they only lead to rigid shift of the $n_s - V_G$ characteristics.

We also investigated the influence of atomistic description of the impurity atoms in the semiconductor on the energy spectrum in the dot and on the shape of the wavefunctions. As an example, in the top left panel of Fig. 5 we show the wavefunction for the fifth eigenstate when atomistic impurity distribution was used.
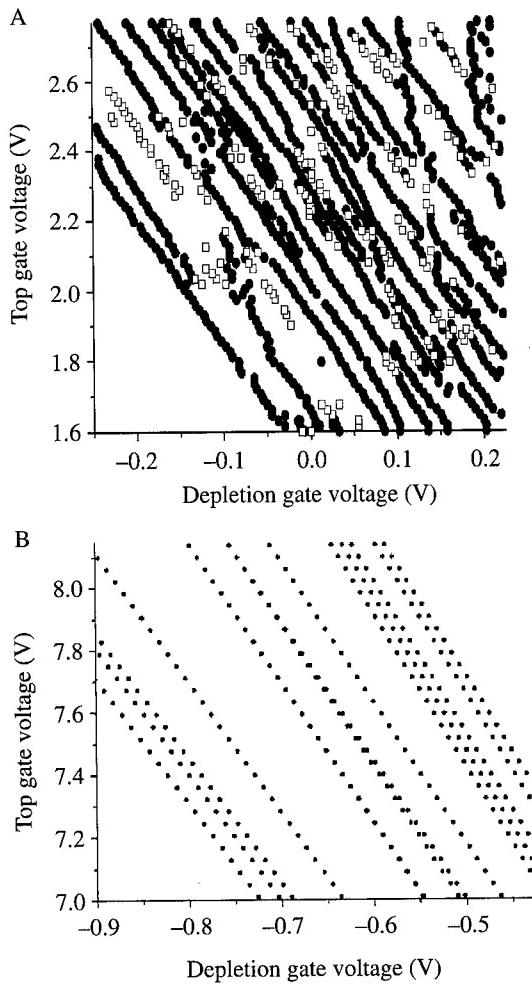
Fig. 4. A, Experimentally measured conductance peaks. B, Conductance peaks derived from simulations of the eigenvalue spectrum.

On the top right panel, the magnitude squared of the expansion coefficients ($|C_n|^2$) versus the eigenvalue number for this particular state is shown. We use as a complete orthonormal basis set the wavefunctions calculated when using uniform impurity distribution. It is quite clear that atomistic impurity distribution, which introduces random fluctuations in the confining potential, leads to mixing of the states. In the bottom panel of Fig. 5 we show the basis functions corresponding to the fourth, fifth and sixth eigenstate which contribute to the mixed state shown in the top left panel.

In conclusion, we have presented simulation results for the energy level spectrum in quantum dots operating in the tunneling regime. We also identified the experimentally observed conductance peaks with energy level in the dot crossing the Fermi level. Examining the influence of atomistic doping description we found out that uniform doping distribution leads to well-defined wavefunctions, and that it is not hard to establish one-to-one correspondence between a wavefunction and a given eigenstate. In the discrete impurity distribution case, mode mixing of the wavefunctions was observed.

Electron wavefunction of the fifth eigenstate
for the discrete impurity distribution case
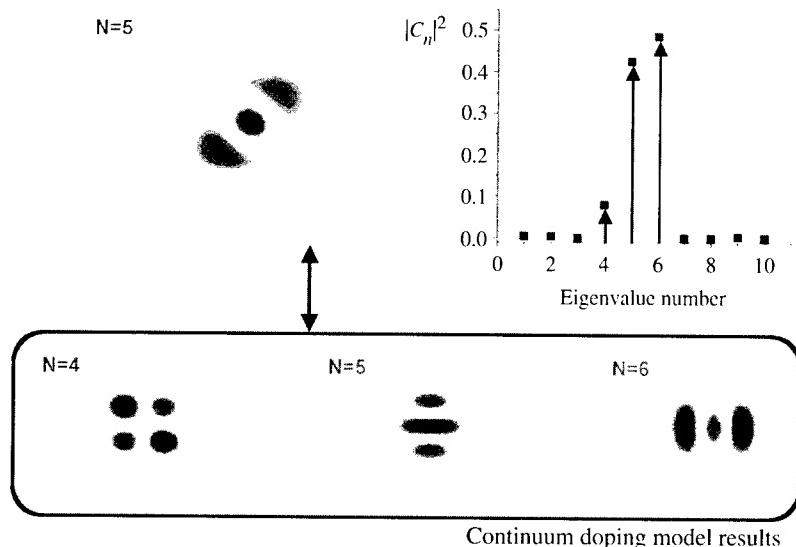
Fig. 5. Wavefunction in the top left, obtained through discrete impurity simulation is a mixed mode consisting of three other modes shown in the lower part of the figure. On the top right panel the magnitude squared of the expansion coefficients ($|C_n|^2$) versus eigenvalue number for this particular state is shown.

References

[1] L. Guo, E. Leobandung, L. Zhuang, and S. Y. Chou, J. Vac. Sci. Technol. **B15**, 2840 (1997).

[2] L. Zhuang, L. Guo, and S. Y. Chou, Appl. Phys. Lett. **72**, 1205 (1998).

[3] J. H. F. Scott-Thomas, S. B. Field, M. A. Kastner, H. I. Smith, and D. A. Antoniadis, Phys. Rev. Lett. **62**, 583 (1989).

[4] H. Ishikuro and T. Hiramoto, Appl. Phys. Lett. **71**, 3691 (1997).

[5] C. Irvine, Z. A. K. Durani, H. Ahmed, and S. Biesemans, Appl. Phys. Lett. **73**, 1113 (1998).

[6] M. Khoury, M. J. Rack, A. Gunther, and D. K. Ferry, Appl. Phys. Lett. **74**, 1576 (1999).

Ultrathin gate dielectrics for silicon nanodevices

M. HIROSE, W. MIZUBAYASHI, KHAIRURRIJAL, M. IKEDA, H. MURAKAMI, A. KOHNO, K. SHIBAHARA, S. MIYAZAKI

Department of Electrical Engineering, Hiroshima University, Higashi-Hiroshima 739-8527, Japan

This paper reviews recent progress in structural and electronic characterizations of ultrathin SiO_2 thermally grown on Si(100) surfaces and applications of such nanometer-thick gate oxides to advanced MOSFETs and quantum-dot MOS memory devices. Based on an accurate energy band profile determined for the n^+-poly-Si/SiO_2/Si(100) system, the measured tunnel current through ultrathin gate oxides has been quantitatively explained by theory. From the detailed analysis of MOSFET characteristics, the scaling limit of gate oxide thickness is found to be 0.8 nm. Novel MOSFETs with a silicon quantum-dot floating gate embedded in the gate oxide have indicated the multiple-step electron injection to the dot, being interpreted in terms of Coulombic interaction among charged dots.

© 2000 Academic Press

Key words: ultrathin SiO_2, MOSFET, gate oxide.

1. Introduction

The high-performance n-MOSFETs with 1.3–1.5 nm thick gate oxides and a large transconductance (gm $= 1.07$ S mm^{-1}) for a 90 nm gate length n-MOSFET with an acceptable level of the gate leakage have been demonstrated [1, 2]. The energy bandgap of 1.7–5.2 nm thick SiO_2 is maintained at 9.85 eV [3], indicating that the intrinsic nature of ultrathin SiO_2 is basically identical to that of bulk oxide although the density of the oxide layer near the SiO_2/Si interface is higher than that of the bulk SiO_2 [4, 5]. The thickness uniformity for 1–5 nm oxides on Si wafers is extremely good because the thermal oxidation of silicon proceeds through a layer-by-layer process on an atomic scale as directly demonstrated by scanning reflection electron microscopy [6].

Quantitative modeling of tunnel leakage current through sub 3 nm oxides is particularly important for future circuit design in which the gate oxide resistance due to direct tunneling current should be taken into account [7]. In the analysis of tunnel current we encounter three crucial problems as follows. (1) An appropriate theoretical framework must be chosen on the basis of WKB approximation [8], multiple scattering theory [9], or transverse-resonant method [10, 11]. (2) The effective mass of tunneling electron must be properly given although the reported values exhibit the broad distribution from 0.29 to $0.5m_0$. (3) Relevant oxide thickness must be determined by ellipsometry, $C-V$ analysis or X-ray photoelectron spectroscopy (XPS). The oxide thickness determined by high-resolution electron microscope (HREM) is higher by 0.2–0.5 nm than the ellipsometric value for 1–2 nm thick SiO_2 [2]. In determining the oxide thickness from $C-V$ curves, poly-Si depletion and quantization effect in the inversion or accumulation layer must be carefully taken into account [10, 11]. Recently it was found that the statistical distribution of direct tunnel leakage current
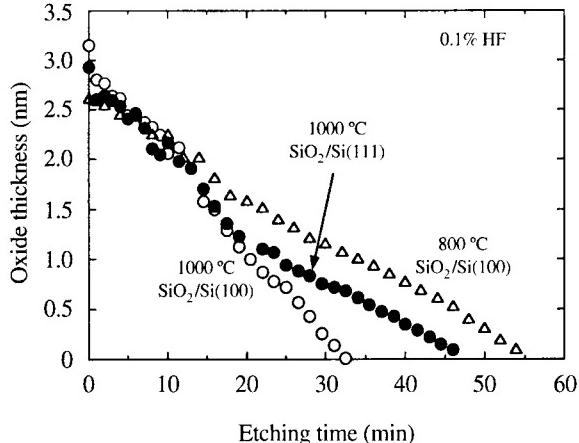
Fig. 1. Oxide thickness as a function of etching time in a 0.1% HF solution. The oxides grown on Si(111) and Si(100) were simultaneously etched in the same HF solution. The oxide thickness uniformity was confirmed by AFM at representative steps of oxide stripping.

through ultrathin gate oxides induces significant fluctuations in the threshold voltage and the transconductance of MOSFETs when gate tunneling resistance becomes comparable to gate poly-Si resistance [12]. A self-assembled silicon quantum-dot layer formed on the ultrathin SiO_2/Si(100) substrate has been used as a floating gate in MOSFETs and the unique memory characteristics has been demonstrated [13]. In this paper, recent studies on structural and electronic properties of ultrathin gate oxides and the application of such thin oxides to advanced silicon nanodevices are reviewed.

2. Growth and structure of ultrathin oxides

The measured oxide thickness on a silicon wafer is extremely uniform as reported in a previous paper, where 64 capacitors on a 200 mm wafer with 3 nm thick, thermally grown oxides yielded the thickness variation of ±0.014 nm [11]. This is basically due to the fact that atomic-scale, layer-by-layer oxidation controls the growth process. Scanning reflection electron microscopy (SREM) has been employed to reveal the initial oxidation of Si(100) surfaces. The periodic reversal of the SREM contrast during oxidation directly showed the layer-by-layer oxidation while interfacial step structure was maintained [6]. The existence of a dense ($\sim$2.4 gm cm^{-3}), thin ($\sim$1 nm) layer at the SiO_2/Si(100) interface has been demonstrated by a high-accuracy difference X-ray reflectively method [4]. Note that the oxide density near the interface is higher than that of bulk SiO_2(2.35–2.36 gm cm^{-3}) as well as bulk Si(2.33 gm cm^{-3}). Recently consistent results have also been reported by another group [5]. Si–O–Si bonds in the interfacial layer (1–1.5 nm thick) with the higher density should be compressively strained. The extent of built-in strain can be detected by an infrared absorption band due to the longitudinal optical (LO) phonon mode originating from Si–O–Si lattice vibration. A distinct redshift of the LO phonon peak from 1250 to $\sim$1210 cm^{-1} is observed within 2 nm from the interface. According to a central force model, the observed redshift of about 30 cm^{-1} for the LO phonon peak corresponds to about a 7.7° reduction of the Si–O–Si average bond angle which is considered to be 144° for the relaxed SiO_2 network [14]. The presence of different levels of compressive, built-in stress in the interface is also suggested by the fact that the etch rate for 800 °C SiO_2 on Si(100) and 1000 °C SiO_2 on Si(111) by dilute HF is significantly reduced in the thickness range below 1.5 nm while the rate for 1000 °C SiO_2 on Si(100), which shows the smaller redshift of the LO phonon peak, is nearly constant as shown in Fig. 1.
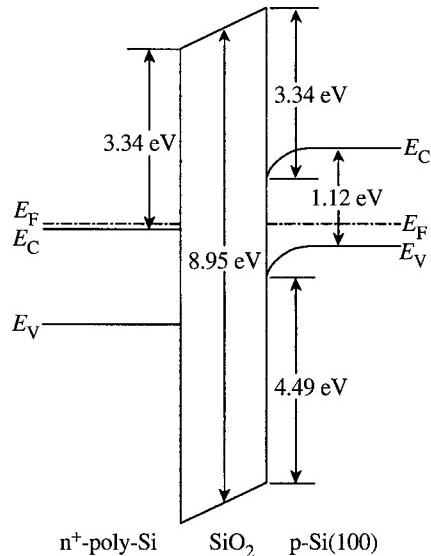
Fig. 2. Energy band profile of an n^+-poly-Si/SiO$_2$/p-Si(100) structure for dry oxides in the thickness range 1.6–40 nm. The wet oxides provide a slightly larger (0.05 eV) electron barrier height. Also, the electron barrier height at the SiO$_2$/Si(111) interface is larger by 0.13 eV than the value indicated in the figure.

3. Electronic states in ultrathin oxides and the SiO$_2$/Si interface

For the purpose of calculating the tunnel current the electron barrier height at the SiO$_2$/Si interface has to be known and the tunnel electron effective mass must be given. Figure 2 represents the energy band profile for an n^+-poly-Si gate/ultrathin SiO$_2$/p-Si(100) structure. Note that the barrier height at the poly-Si/SiO$_2$ interface is tentatively set equal to the value at the SiO$_2$/Si(100) interface. For obtaining the barrier height we have determined simultaneously the oxide bandgap and the valence band alignment at the SiO$_2$/Si interface for 0.8–5.2 nm thick oxides by using high resolution X-ray photoelectron spectroscopy [15]. The oxide bandgap has been obtained from the energy loss spectrum of the O$_{1s}$ core level photoelectron because the photoexcited electrons in SiO$_2$ suffer inelastic losses such as plasmons and band-to-band excitation whose onset appears at 8.95 eV for oxides thicker than 1.7 nm, as shown in Fig. 3. For oxides thinner than 1.2 nm, the onset of the bandgap excitation becomes less clear and a significant energy loss yield emerges in the subgap energy region mainly because the energy loss due to the interface suboxide states contributes to the 5–9 eV signal. Further, the integrated intensity of enhanced O$_{1s}$ energy loss signal over 4–10 eV is not proportional to the integrated Si$_{2p}$ suboxide signal intensity for oxides thinner than 0.8 nm. This might arise from the compressive, built-in stress in the interfacial oxide layer because the stress changes the local density of states in SiO$_2$ and induces the localized band tail states below the conduction band [16].

4. Modeling of tunneling current

The effective mass for tunneling electron has been empirically determined as a kind of fitting parameter based in most cases on the WKB approximation. The reported values of electron effective mass for the direct tunnel current regime are in the range 0.29–0.32m_0 for the WKB approximation with a parabolic E–k relationship [9, 17–19]. On the other hand, the tunnel current more accurately calculated by the multiple scattering theory has been fitted to measured current with the effective mass of 0.36m_0 for oxides thicker than 3 nm [9]. We have also compared measured tunnel current with the multiple scattering theory for sub
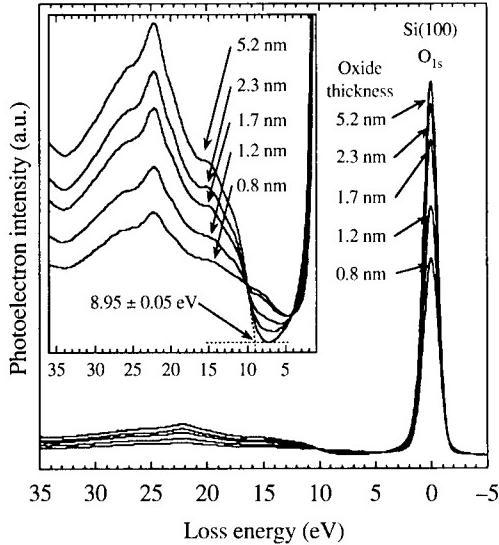
Fig. 3. O_{1s} photoelectron and its energy loss spectra for oxides grown on Si(100) at 1000 °C.

5 nm oxides, and the best fit has been obtained for $m^* = 0.35m_0$ as shown in Fig. 4 [20]. In the calculation of tunnel current by the multiple scattering theory both forward and backward tunneling are taken into account, whereas there is significant discrepancy between the measured and calculated current at the oxide voltages below 1 V. This arises from the difficulty in predicting exact electron density at the n^+-poly-Si/SiO$_2$ interface which significantly varies at low negative gate voltages. It should be noted that at the oxide voltages above 3.34 V, which is equal to the electron barrier height, Fowler–Nordheim tunneling current oscillation is also quantitatively reproduced as understood from I–V curves for 3.75–4.82 nm oxides in Fig. 4, where the oxide conduction band effective mass of $0.60m_0$ is used together with the direct tunneling effective mass of $0.35m_0$. To improve the discrepancy between the calculated current and measured one at $V_{ox} < 1$ V, a new analytic model has been developed [21]. Using the nonparabolic E–k dispersion relationship for the oxide bandgap, the calculated direct tunnel current densities through the gate oxides are compared with the measured data over the oxide thickness range of $T_{ox} = 1.65$–3.9 nm, as shown in Fig. 5. It is clearly shown that the calculated current densities fit well to the measured data over the entire oxide voltage range. It is found that the effective mass increases with decreasing T_{ox}, presumably due to compressive stress in the oxide layer near the SiO$_2$/Si(100) interface. The effective mass for $V_{ox} > 1.5$ V is about a factor of 0.8 times smaller than that of $V_{ox} < 1.5$ V. The reason for this could be qualitatively explained as follow. For low V_{ox} direct tunnel occurs from the Fermi level of the n^+-poly-Si gate to the Si(100) indirect conduction band, while for high V_{ox} the final states are the Si(100) direct conduction band whose effective mass is smaller than that for the Si(100) indirect conduction band [22, 23]. Since electron waves in the n^+-poly-Si, the substrate Si(100) conduction bands and the oxide bandgap mutually interact through the ultrathin oxide layers, it is likely that the effective mass of the final states influences the value of obtained effective mass.

5. Ultrathin gate oxide MOSFETs

The oxide thickness scaling is an effective way for device performance improvement, while it leads to increases in gate tunnel leakage current I_g. We have fabricated n-MOSFETs with 1.2–2.8 nm thick gate oxides to investigate the influence of direct tunnel leakage current on the dc characteristics of MOSFETs [12].
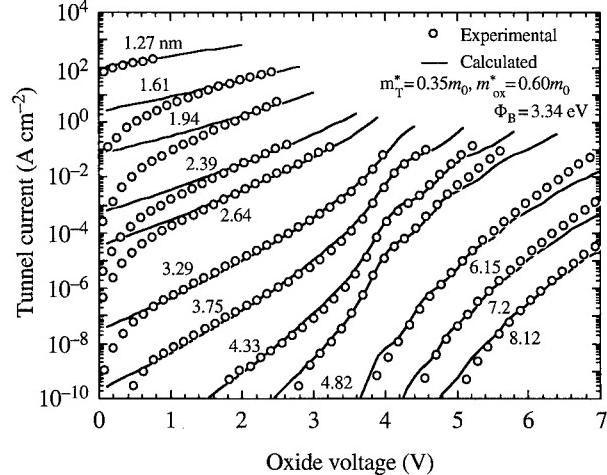
Fig. 4. Tunnel current through ultrathin oxides for n^+-poly-Si gate MOS diodes measured as a function of voltage across the oxide. The tunneling effective mass $m_T^* = 0.35m_0$ and the oxide conduction band effective mass $m_{ox}^* = 0.60m_0$ are used in the calculation.
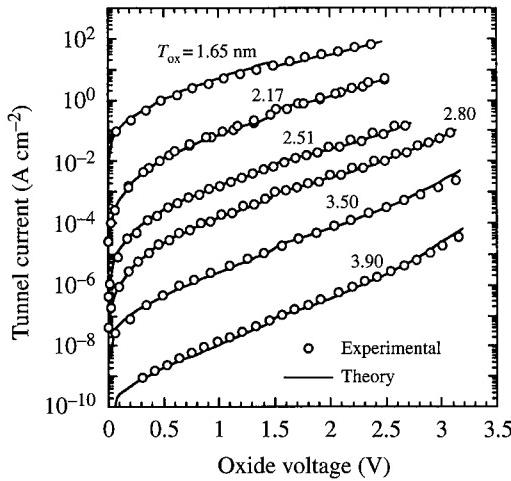

Fig. 5. Measured and calculated direct tunnel current densities through 1.65–3.90 nm thick oxides as a function of oxide voltage.

As shown in Fig. 6, the threshold voltage V_{th} for 1.2 nm thick gate oxide MOSFETs significantly increases in proportion to I_g while V_{th} remains nearly constant regardless of the I_g distribution for 2.0 nm thick gate oxide MOSFETs. The result is explained by a simple model as follows. The gate current flows into a channel via the gate poly-Si and gate oxide as illustrated in Fig. 7A. When the gate oxide resistance R_{SiO_2} is lowered by T_{ox} scaling, the voltage drop due to the gate poly-Si resistance R_{poly} can no longer be neglected. As a result, V_{th} apparently increases by the product of R_{poly} and I_g as explained in Fig. 7B. In Fig. 8, R_{SiO_2} as defined by $\partial V_{ox}/\partial I_g$ at $V_g = 1.5$ V where I_g is the calculated tunnel current by multiple scattering theory [20] is plotted as a function of oxide thickness together with experimental data. As references, resistances of poly-Si gate R_{poly} (this work) and salicided gate $R_{salicide}$ and metal gate R_{metal} are indicated in the figure. Considering that $R_{salicide}$ for a $0.1 \times 10 \ \mu m^2$ gate area is 1000 Ω, R_{SiO_2} becomes comparable to $R_{salicide}$ when T_{ox} is
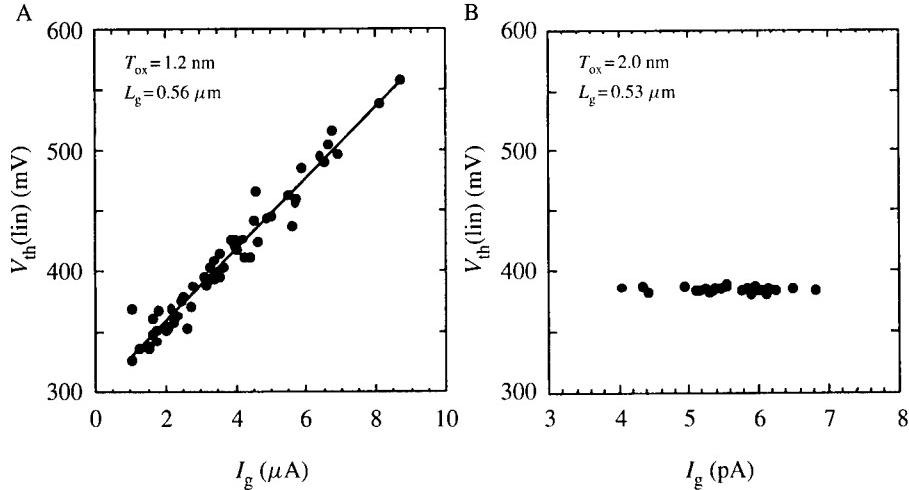
Fig. 6. Relationship between linear mode V_{th} and I_g distribution at $V_g = 1.5$ V or at V_{th} for A, 1.2 nm and B, 2.0 nm gate oxide MOSFETs.
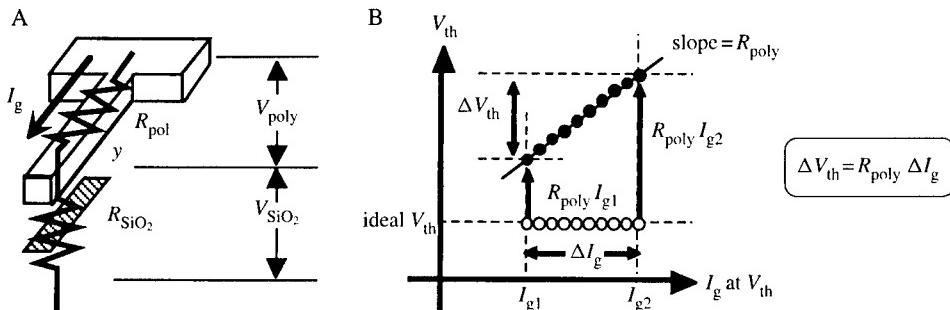

Fig. 7. Schematic explanation of V_{th} fluctuation induced by I_g distribution. Equivalent circuit of I_g path A, and V_{th} shift or fluctuation ΔV_{th} caused by I_g variation B.

decreased to 0.8 nm. Namely, even if the gate electrode resistance is reduced by using salicide, V_{th} and G_m fluctuations will emerge when the T_{ox} is reduced to 0.8 nm.

6. Silicon quantum-dot floating gate MOS structures

Single-crystalline Si quantum dots (QDs) were self-assembled on 3.5 nm thick SiO_2 by LPCVD of pure SiH_4 at 580 °C [24]. After the first-layer Si dot array formation a $\sim$1 nm thick oxide layer was grown and the second dot array was deposited under the same conditions. The average dot height and diameter evaluated by AFM were 5 and 10 nm, respectively. The total dot density was $\sim$6 × 10^{11} cm^{-2} for MOSFETs and $\sim$5 × 10^{11} cm^{-2} for MOS capacitors. The surface of the second Si dot layer was covered by $\sim$1 nm thick oxide. Subsequently, a 3.3 nm thick amorphous Si layer was grown over the Si QD layer by LPCVD at 440 °C and fully oxidized in dry O_2 at 1000 °C to fabricate a 7.5 nm thick control oxide. No significant change in the surface morphology was observed in this oxide layer fabrication process. Finally, 300 nm thick n$^+$-poly-Si gates were grown.
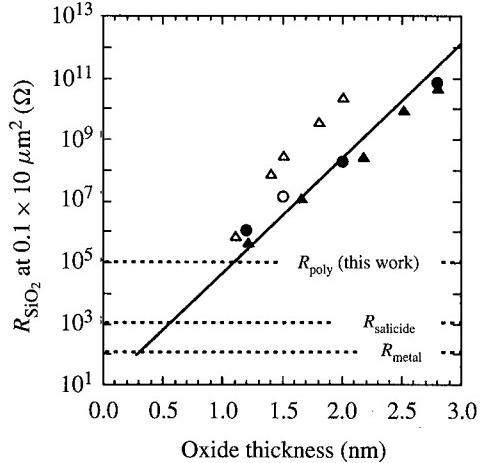
Fig. 8. Relationship between oxide resistance R_{SiO_2} and oxide thickness. R_{SiO_2} was obtained by $E_{ox}T_{ox}$ divided by I_g while R_{SiO_2} for Ref. [1] was evaluated by $V_g(= 1.5$ V) divided by I_g. — Calculated by MST, $E_{ox} = 5.0$ MV cm^{-1}. ●, this work MOSFET, $E_{ox} = 5.0$ MV cm^{-1}; ▲, this work MOS diode, $E_{ox} = 5.0$ MV cm^{-1}; ○, MOS diode, $V_g = 1.5$ V [1]; △, MOS diode, $E_{ox} = 5.0$ MV cm^{-1} [2].

## 6.1. Memory operation of Si quantum-dot floating gate MOSFETs

Drain current versus gate voltage characteristics of a Si QD floating gate MOSFETs are shown in Fig. 9. The hysteresis arises from the charging of a Si QD floating gate. The bumps around 0.3 and 0.8 V are caused by the electron charging to the Si QD floating gate. Temporal change in the drain current for a single-pulse gate bias was measured, as shown in the inset of Fig. 10. The drain current drops to the minimum value at t_1 and gradually increases to the final steady state value. This indicates that a fraction of electrons retained in the dots at a positive gate pulse is released to the channel at zero gate bias. The minimum drain current versus the gate voltage pulse height is shown in Fig. 10. The stepwise changes in the minimum drain current around 0.2, 0.4 and 0.8 V suggest that the multilevel electron charging to Si QDs occurs by applying a positive gate pulse. It is likely that the electron distribution in the dot layer is rearranged when total injected memory charge density is increased because Coulombic interaction among charged dots induces lateral tunneling.

6.2. Electron charging to Si quantum-dot floating gate MOS capacitors

Electron charging characteristics of MOS capacitors with a QD floating gate have also been investigated [13]. The measured $C-V$ characteristics are shown in Fig. 11. When the gate bias was swept over the range of -4 to $+6$ V, the $C-V$ curve showed a unique hysteresis which arises from electron charging to the dots through the 3.5 nm thick bottom oxide. The steady state $C-V$ curves after electron charging at 6 V or after discharging at -4 V were obtained from the transient capacitance characteristics. The charged and discharged states are maintained longer than 10^3 s at gate voltages between -0.7 and $+0.7$ V. Gate voltages higher than $+0.7$ V the steady state $C-V$ curve after charging is identical to that after discharging. This is because the tunnel injection current from the substrate to the dots balances the dots-to-substrate and/ or the dots-to-gate emission current, and the number of stored electrons per dot determines the capacitance value. On the other hand, at gate voltages lower than -0.7 V, electrons in the Si dots are completely released and the device stays in the discharged state or the memory charge is erased.
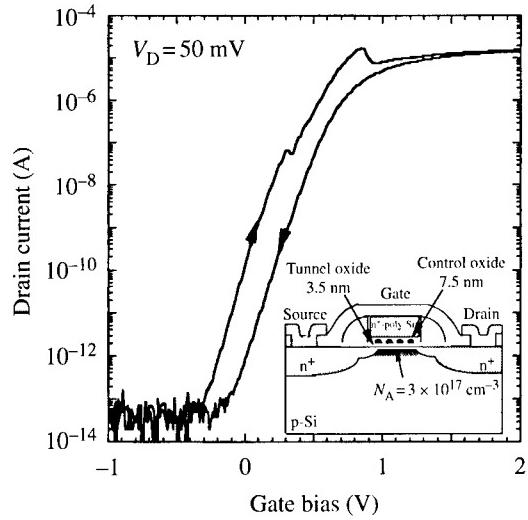
Fig. 9. Drain current–gate voltage characteristics of a Si QD floating gate MOSFET. The gate length and width are 0.5 μm and 10 μm, respectively.
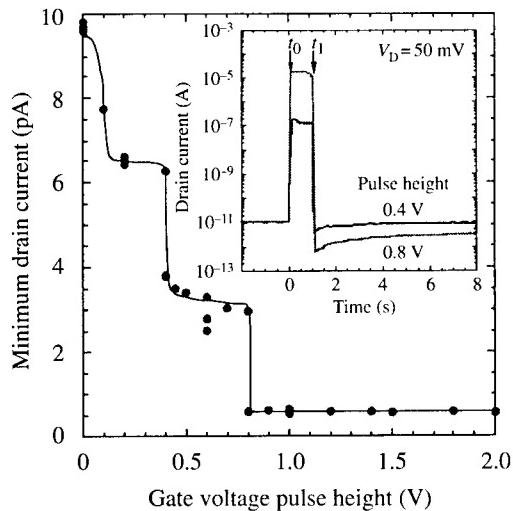
Fig. 10. Minimum drain current as a function of the gate voltage.

It is essential for control of the memory operation to understand the relationship between the flat-band voltage shift and the stored electron number per dot. To evaluate the total charge stably stored in QDs around zero bias, the external gate bias was kept at 0 V after writing operation and switched to −1.2 V (erased state), and the discharge current was measured. By integrating the transient current, the total dot charge Q_{dot} is obtained to be 2.6×10^{-10} C, which corresponds to the number of retained electron per dot being ~1.3. Thus, the result indicates that the storage of approximately one electron per dot causes the flat-band voltage shift $\Delta V_{FB} = 0.7$ V.
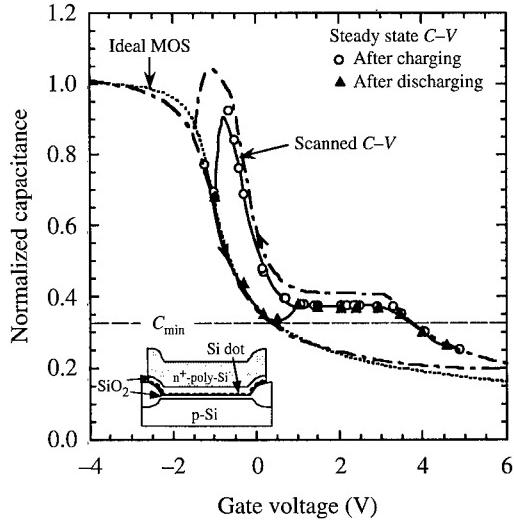
**Fig. 11.** *C–V* characteristics of a Si QD floating gate MOS capacitor. The steady state *C–V* curves were obtained from the transient response under constant bias.

7. Sub 30 nm gate length MOSFETs

Fabrication procedure of a 30 nm gate length n-MOSFET is as follows. The gate resist pattern was transferred to an SiO_2 layer on n^+-poly-Si by RIE. Subsequently the SiO_2 mask was shrunk by chemical etching in 0.5% HF. Precise control of the gate mask shrinking is possible by using dilute HF etch, as shown in Fig. 12. For n^+-poly-Si gate formation, ECR etching was employed in a Cl_2/N_2 gas mixture at 5 m Torr with a self-bias of −80 V. An etching gas for n^+-poly-Si gate etch was changed from a Cl_2/N_2 mixture to Cl_2/O_2 just before the etch end point in order to improve the etch selectivity of poly-Si to SiO_2 as high as 400. The cross-sectional SEM image of the etched gate pattern is shown in Fig. 13, where the gate length is 30 nm. The source/drain extension was formed by Sb^+ ion implantation at 10 keV which realizes a junction depth of 20 nm. The measured drain current versus drain voltage characteristics of a 26 nm gate length MOSFET with a gate oxide thickness of 1.86 nm are shown in Fig. 14. The gate leakage current is not significant even at $V_g = 1.5$ V. The measured transconductance of 430 mS mm^{-1} is about half of the simulated result because of parasitic resistance induced by a long distance (4 μm) from the channel region to the contact holes.

8. Summary

In summary we have shown that layer-by-layer silicon oxidation mechanism helps formation of extremely uniform, nanometer-thick gate oxides for advanced MOSFETs. Measured tunnel current through ultrathin oxides agrees well with theory without taking into account SiO_2 thickness nonuniformity, in consistency with the abrupt, nearly defect-free SiO_2/Si interface structure. Good performance of 26–100 nm gate length MOSFETs with 1.2–2.0 nm thick gate oxides has been demonstrated. Silicon quantum-dot floating gate MOS memory devices have been fabricated. It is shown that approximately one electron per quantum dot is stably stored during memory retention period and electron charging into the dots occurs through a multiple-step process due to Coulombic interaction among charged dots.
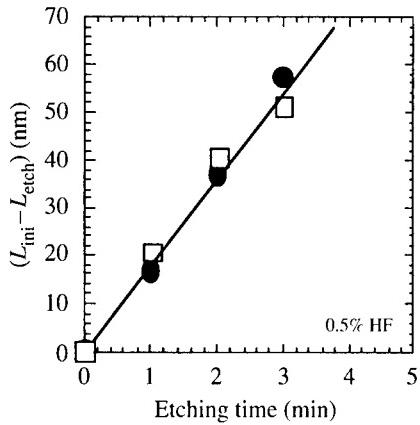
Fig. 12. SiO_2 gate mask shrinking rate versus etching time. Here, L_{ini} is the initial gate length and L_{etch} is the gate length after shrinking. Solid circles and open squares refer to different runs.
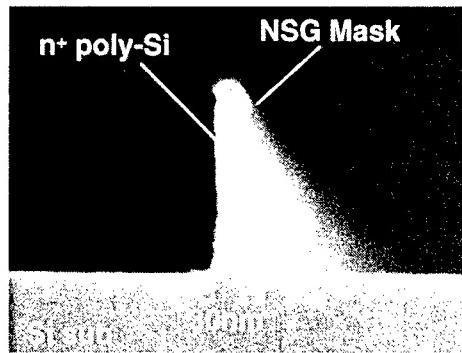
Fig. 13. Cross-sectional SEM image for 30 nm gate pattern.
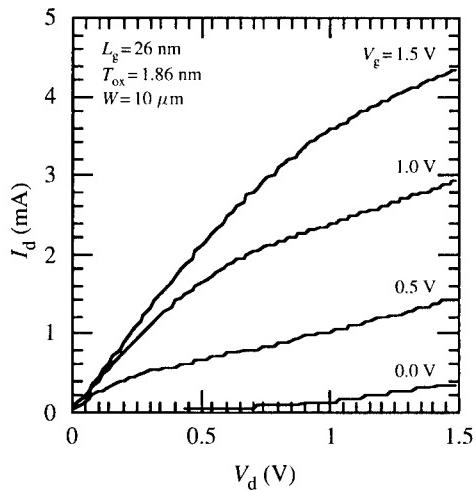
Fig. 14. I_d–V_d curves for a 26 nm gate length MOSFET.

References

[1] H. S. Momose, M. Ono, T. Yoshitomi, T. Ohguro, S. Nakamura, M. Saito, and H. Iwai, Tech. Dig. IEDM, 593 (1994).

[2] G. Timp *et al.*, Tech. Dig. IEDM, 930 (1997).

[3] S. Miyazaki, H. Nishimura, M. Fukuda, L. Ley, and J. Ristein, Appl. Surf. Sci. **113/114**, 585 (1997).

[4] N. Awaji, S. Ohkubo, T. Nakanishi, Y. Sugita, K. Takasaki, and S. Komiya, Jpn. J. Appl. Phys. **35**, L67 (1996).

[5] S. D. Kosowsky, P. S. Pershan, K. S. Krisch, J. Bevk, M. L. Green, D. Brasen, L. C. Feldman, and P. K. Roy, Appl. Phys. Lett. **70**, 3119 (1997).

[6] H. Watanabe, K. Fujita, T. Kawamura, and M. Ichikawa, Phys. Rev. Lett. **80**, 345 (1998).

[7] B. E. Weir *et al.*, Tech. Dig. IEDM, 73 (1997).

[8] T. Yoshida, M. Imafuku, J. L. Alay, S. Miyazaki, and H. Hirose, Jpn. J. Appl. Phys. **34**, L903 (1996).

[9] S. Nagano, M. Tsukiji, K Ando, E. Hasegawa, and A. Ishitani, J. Appl. Phys. **75**, 3530 (1994).

[10] S.-H. Lo, D. A. Buchanan, Y. Taur, L.-K. Han, and E. Wu, Symposium on VLSI Technol., Digest of Technical papers, 149 (1997).

[11] D. A. Buchanan and S.-H. Lo, *The Physics and Chemistry of SiO$_2$ and Si–SiO$_2$ Interface-3*, edited by H. Z. Massoud, E. H. Poindexter, and C. R. Helms (The Electrochemical Society, Inc., Pennington, NJ, 1996) p. 3.

[12] M. Koh *et al.*, Tech. Dig. IEDM, 919 (1998).

[13] A. Kohno, H. Murakami, M. Ikeda, H. Nishiyama, S. Miyazaki, and M. Hirose, *Extended Abstracts of Intern. Solid State Devices and Materials* (Business Center for Academic Societies, Japan, Tokyo, 1998) p. 174

[14] T. Yamazaki, S. Miyazaki, C. H. Bjorkman, M. Fukuda, and M. Hirose, Mat. Res. Soc. Symp. Proc. **318**, 419 (1994).

[15] J. L. Alay and M. Hirose, J. Appl. Phys. **81**, 1606 (1997).

[16] R. B. Laughlin, J. D. Jannopoulos, and D. J. Chandi, *The Physics of SiO$_2$ and its Interfaces*, edited by S. T. Pantelides (Pergamon Press, NY, 1978) p. 321.

[17] M. Depas, R. L. Van Meirhaeghe, W. H. Laflere, and F. Cardon, Microelectron. Eng. **22**, 61 (1993).

[18] M. Hiroshima, T. Yasaka, S. Miyazaki, and M. Hirose, Jpn. J. Appl. Phys **33**, 395 (1994).

[19] B. Brar, G. D. Wilk, and A. C. Seabough, Appl. Phys. Lett. **69**, 2728 (1996).

[20] M. Fukuda, W. Mizubayashi, S. Miyazaki, and M. Hirose, Jpn. J. Appl. Phys **37**, L1534 (1998).

[21] Khairurrijal, W. Mizubayashi, S. Miyazaki, and M. Hirose, *Workshop on Silicon Nanofabrication and Nanodevices* (Core Research for Evolutional Science and Technology, Tokyo, 1997) p. 37.

[22] J. Maserjian and G. P. Petersson, Appl. Phys. Lett. **25**, 50 (1974); 50 (1974).

[23] S. M. Sze, *Physics of Semiconductor Devices* (John Wiley, New York, 1981) 2nd edn, p. 13.

[24] M. Fukuda, K. Nakagawa, S. Miyazaki, and M. Hirose, Appl. Phys. Lett. **70**, 2291 (1997).

Oxide reliability improvement controlling microstructures of substrate/oxide and oxide/gate interfaces

JIRO YUGAMI[†]

Central Research Laboratory, Hitachi., Ltd., Kokubunji, Tokyo 185, Japan

For future ULSIs, the oxide reliability problem is a key issue to realize low-power, high-speed devices whilst retaining its reliability. In the MOSFET structure, a gate oxide consists of the substrate/oxide interface, oxide and oxide/gate interface. Therefore, to improve oxide reliability, it is important to control these three component structures individually. In this report, I will describe experiments to control structures of the above two interfaces using: (1) a combination of a closed wet cleaning system and a load-lock oxidation system and (2) an ultra-thin film deposition CVD technique. By controlling these structures, the oxide reliability was improved. Moreover, the interface structure should be carefully controlled in high-k gate dielectrics in future devices.

© 2000 Academic Press

Key words: MOSFET, oxide, interface structure.

1. Introduction

Oxide failure is one of the most serious problems anticipated in giga-scale ULSIs. As integration generation improves, the total gate area expands and the electric field in the oxide increases [1]. The gate oxide reliability decreases as the gate area increases due to weak spots in the oxide [2,3], and the high electric field accelerates the oxide's breakdown. Thus, we predict that the failure rate of the gate oxide will increase drastically in giga-scale ULSIs.

In the MOSFET structure, the gate oxide consists of three components; (1) the substrate/oxide interface, (2) oxide and (3) the oxide/gate interface. Therefore, to improve oxide reliability, it is important to control these three component structures individually.

To control the substrate/oxide interface structure, we used a closed wet cleaning (CWC) system and a load-lock oxidation system. Using these systems, we controlled the Si surface just before the oxidation process and the moisture during oxide formation. In this way it may be possible to obtain almost completely ideal oxide formation keeping flat and contamination-free substrate/oxide interface.

To control the oxide/gate interface structure, we used an ultra-thin film deposition CVD technique. It was found that nm-grain-sized poly-Si could be formed by annealing the nm-thick a-Si films [15]. In this case the grain size decreases by reducing the a-Si film thickness. Thus, the oxide/gate interface was controlled by nm-grain structures. In this report, I will describe experiments to control structures of the above two interfaces using: (1) a combination of a closed wet cleaning system and a load-lock oxidation system and

---

[†]E-mail: yugami@crl.hitachi.co.jp
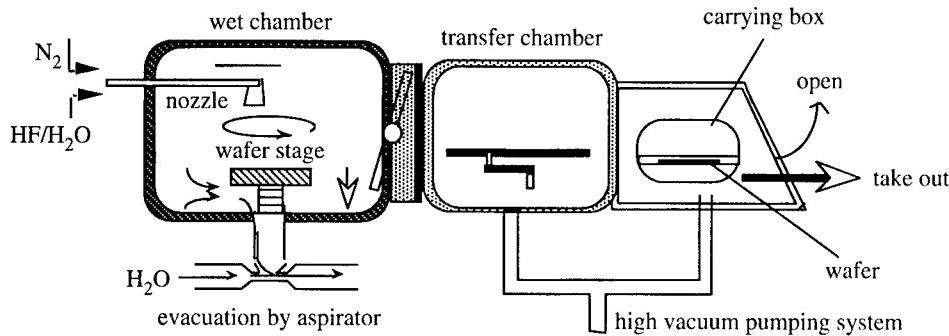
Fig. 1. The closed wet cleaning (CWC) system.

(2) an ultra-thin film deposition CVD technique. By controlling these structures, the oxide reliability was improved.

2. Experiments

2.1. Closed wet cleaning system

To control the Si surface just before the oxidation process, we developed a closed wet cleaning (CWC) system (Fig. 1). First, the Si wafer is settled in a carrying box in the clean room. Next, this box is settled in a loading chamber to displace clean room air with pure N_2 using a pumping system. Then, the Si wafer is transferred to the wet cleaning chamber. This wet cleaning chamber allows us to clean the Si wafer with HF/H_2O solution without exposure to clean room air. After treatment in HF solution, the Si wafer is dried with hot N_2. In this operation, the HF containing gas is displaced with pure N_2 using the aspiration evacuate system. Finally, the cleaned Si wafer is returned to the carrying box that is filled with pure N_2 gas.

2.2. Load-lock oxidation system

Si wafers cleaned with the CWC system were carried to the load-lock oxidation system using the SMIF box to avoid exposure to the clean room air. This oxidation system (Fig. 2), allows us to control the wafer ambient moisture in the oxidation step from 3 ppb to 50% using a bubbler and pyrogenic burner. After the oxidation the oxidation chamber is evacuated to less than 10^{-4} Pa. By supplying dry N_2 and H_2O from the bubbler to the oxidation chamber, we can control the unloading atmosphere to keep the moisture from 3 to 1000 ppm. The oxidation sequence used in this work is shown in Fig. 3.

2.3. Nanometer-grain poly interface

To form the nm-grain poly interface, the following process sequence was performed. After gate oxide formation, ultra-thin, 2–10 nm thick, a-Si film was deposited at 425 °C using the LPCVD technique. Next, in order to form a large grain poly-Si, a 200 nm thick a-Si film was deposited at 525 °C. To crystallize these Si films, 900 °C annealing in N_2 ambient was performed for 20 min, as shown in Fig. 4. From the TEM observation, we found that crystallization in the lower layer began at 800 °C while the upper layer crystallizes easily at 600 °C. As a result, a nm-grain/micron-grain double-layer poly-Si gate structure was formed as shown in TEM results at Fig. 4. Additionally, this large grain size and the flat surface of the upper portion of this gate structure can provide the highly reliable inter-poly dielectric films.

Fig. 2. The load-lock oxidation system.

Fig. 3. Oxidation sequences in the load-lock oxidation system.
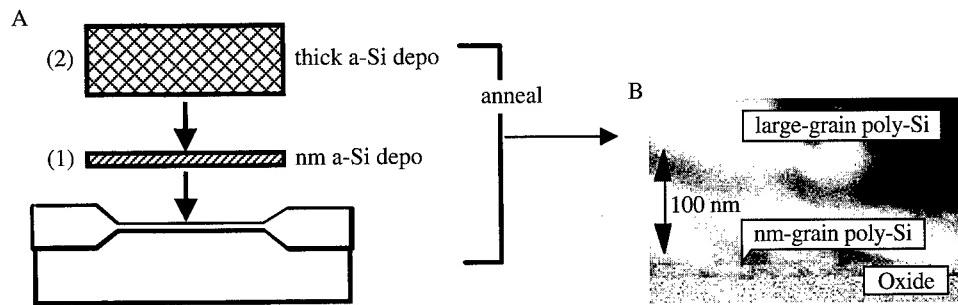
Fig. 4. A, Process sequence for a nm-grain/large-grain double-layer poly-Si gate. B, TEM observation of a nm-grain/large-grain double-layer poly-Si gate.

2.4. Electrical measurement

Using the above technique, we fabricated n^+-poly-Si gate MOS capacitors with various capacitor areas ($10^{-6} \sim 0.5$ cm^2) on a p-type (100) Si substrate. In these capacitors, the gate edge was on the field oxide formed by the LOCOS technique. $I-V$ measurements were performed applying step ramp voltage to the gate electrode. In each voltage step, the gate current was measured after a delay of 1 s. The electric field
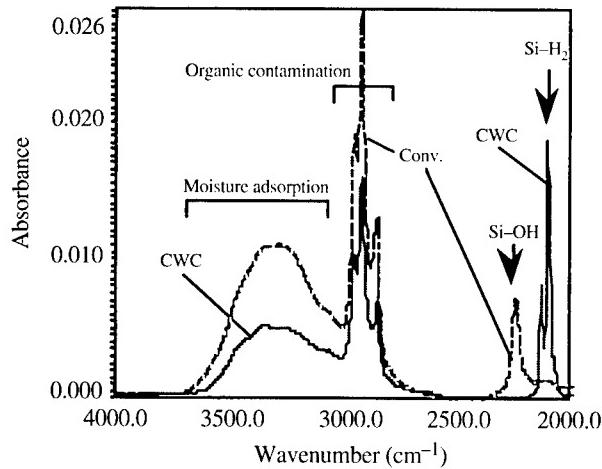
Fig. 5. Comparison of FT-IR/ATR spectra obtained from Si surfaces with different surface treatments. One is after conventional wet station cleaning (Conv.) and the other is after H termination in the closed wet cleaning (CWC) system.
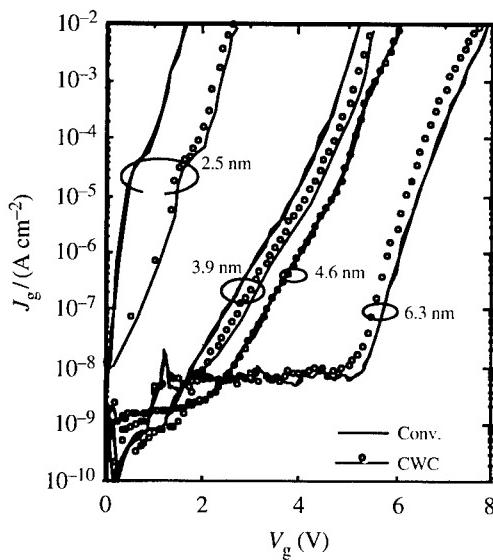
Fig. 6. $J-V$ characteristics of MOS capacitors with thin gate oxides. Gate oxidation processes were performed after conventional pre-cleaning in wet station (Conv.) or H termination with closed wet cleaning system (CWC).

in oxide, E_{ox}, was calculated using $E_{ox} = (V_g - 1.1)/t_{ox}$, where V_g represents the gate voltage and t_{ox} is the oxide thickness measured by ellipsometry and capacitance measurements. The constant voltage TDDB measurement was performed using HP-4142B by applying a pulse voltage to the gate electrode. The pulse width was varied from 1 ms to 10 s automatically, depending on breakdown times. The breakdown was identified by gate current at a low electric field (4 MV cm^{-1}) after each pulse stressing.
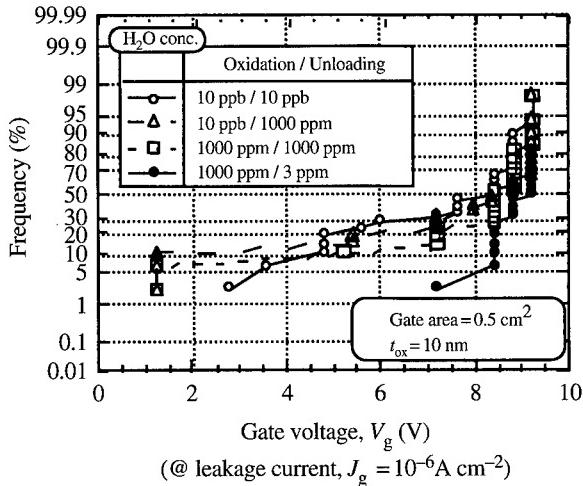
Fig. 7. The breakdown voltage of MOS capacitors dependence on H_2O concentration in oxidation and unloading step.
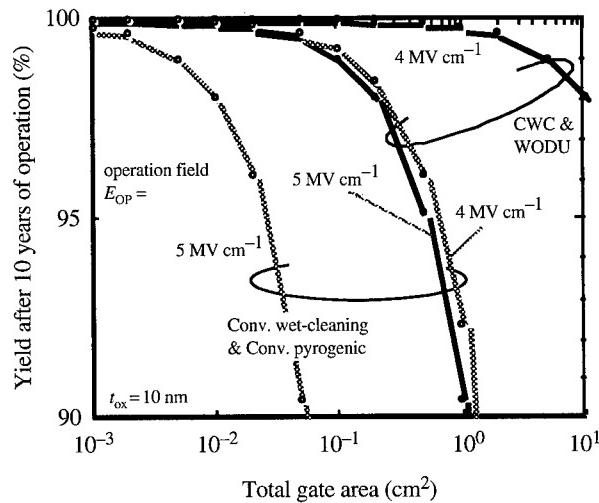
Fig. 8. Predicted gate oxide yield in LSI chip after 10 years of operation under 4 MV cm^{-1} and 5 MV cm^{-1} oxide fields.

3. Results and discussion

3.1. Surface control with the CWC system

To investigate the effectiveness of the CWC system, we analysed the Si surface after HF/H_2O treatment in this system using an *in situ* ATR-FTIR measurement system. After NH_4/H_2O_2 cleaning in a conventional wet station, a p-type Si (100) wafer was loaded in the CWC system using the SMIF box. Then, a Si wafer was treated with 0.5% HF/H_2O solution in a wet chamber followed by hot N_2 blow on the spinning stage. Figure 5 shows the ATR–FTIR spectrum of a Si surface which was treated with 0.5% HF/H_2O solution in the CWC system compared with a Si surface treated in a conventional open-air wet station. In the CWC system a strong Si–H_2 peak is obtained which means that the Si surface is successfully H-terminated without growth of native
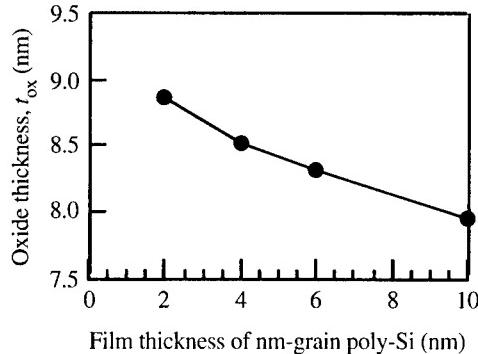
Fig. 9. The effective tunnel oxide thickness, t_{ox}, dependence of nm-grain poly-Si thickness.

oxide. Moreover, the moisture adsorption and organic contamination on the Si surface normally observed in conventional wet pre-cleaned surfaces are greatly reduced. These results suggest that the CWC system is effective for forming a clean (contamination-free and native oxide-free) and atomic-scale flat surface [4].

3.2. Tunneling current reduction by surface control

To investigate the effectiveness of the surface-control technique in the CWC system, I–V characteristics of MOS capacitors with thin gate oxide were measured. In this experiment, a Si–H-terminated Si wafer was transferred to the load-lock oxidation system using the SMIF box. Thus, the surface just before the oxidation process is almost the same as the surface controlled in the CWC system. Then the wafer was oxidized in dry O_2 ambient at 800 °C. Figure 6 shows the results accompanied with MOS capacitors whose gate oxide was formed after conventional wet pre-cleaning in an open-air wet station. The Si–H-terminated Si surface is clearly effective in reducing the direct tunneling current of thin (<4 nm) oxide [5]. To explain this result, we assumed that the direct tunneling current may easily increase with surface contamination or poor-quality native oxide just before the oxidation process, which strongly affects the n^+-poly-Si/SiO$_2$ interface.

3.3. Moisture control in oxidation process

From the results in Fig. 7, a guideline for improving oxide reliability is obtained; i.e. to oxidize the Si wafer in wet ambient and to diminish the moisture ultimately in the unloading step (cooling step) [6]. The effectiveness of H_2O in the oxidation step can be explained by FTIR analysis of the obtained oxide in the case of the ultra-dry unloading condition. The TO peak of the Si–O bond shifts to a higher wavenumber with increasing H_2O concentration in the oxidation step. This result suggests that even the ppm-H_2O in the oxidation step can reduce the oxide film stress [11] that may cause weak spot formation [9]. On the other hand, H_2O in the unloading step may cause degradation of oxide reliability by accelerating the growth of poor-quality oxide at low temperatures [10, 12]. This oxide growth may lead to a poor substrate/oxide interface structure. Thus, ultra-dry unloading is the key to keeping good oxide quality.

3.4. Reliability prediction of gate oxide from CWC system for giga-scale ULSIs

To investigate the impact of the CWC and WODU technique on the reliability of LSIs, the weak spot densities in the oxide formed by the combination of CWC and WODU were analysed [3, 8, 9] based on the TDDB measurement by using MOS capacitor sets with differing areas of $10^{-6} \sim 0.5$ cm^2. Using the
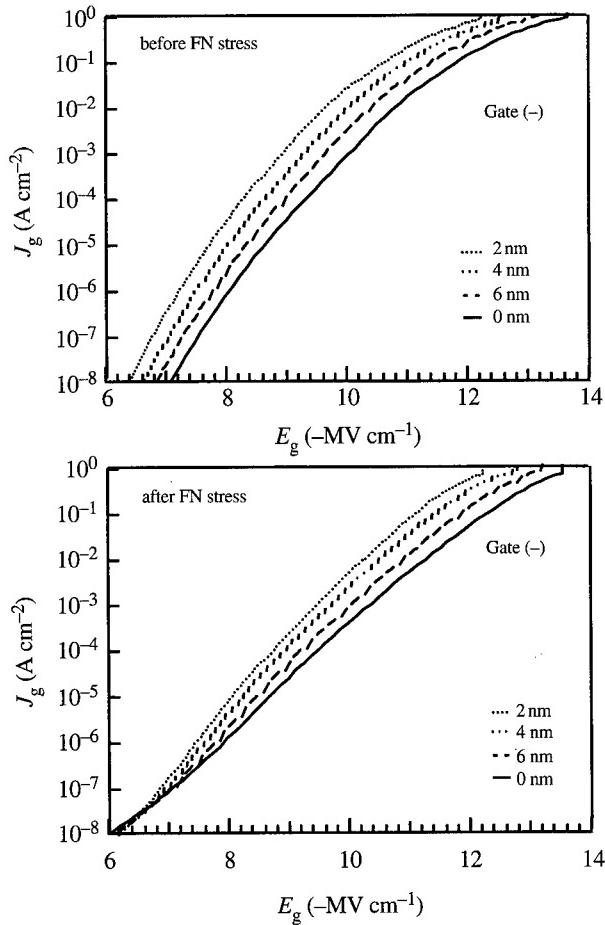
Fig. 10. J–V characteristics of MOS capacitors with a nm-grain double-layer poly-Si gate.

calculated weak spot densities in gate oxide formed by various techniques, we predicted the reliability of the gate oxide from the combination of CWC and WODU as functions of total gate area and operating electric fields. Figure 8 shows the results which indicate that the CWC and WODU technique significantly improved oxide reliability, showing that a 90% yield is expected after 10 years of operation, even when the total gate area is as large as 1 cm^2 and the electric field is as high as 5 MV cm^{-1}.

E. The gate poly-Si structure impacts on oxide reliability

(i) *Film thickness increase.* In the case of a nm-grain poly-Si double-layer gate electrode, the effective tunnel oxide thickness increases with decreasing nm-grain poly-Si layer thickness, as found from the C–V measurement and also from ellipsometry (Fig. 9).

(ii) *Tunnel current increase.* On the other hand, the tunneling current under the same electric field increases in thinner nm-grain layers as shown in Fig. 10, when the gate was negatively biased. This increase of tunnel current occurs when the nm-grain layer thickness is less than 8 nm.

(iii) *Impact on SILC.* The stress-induced leakage current (SILC) after 8 C cm^{-2} injection does not depend
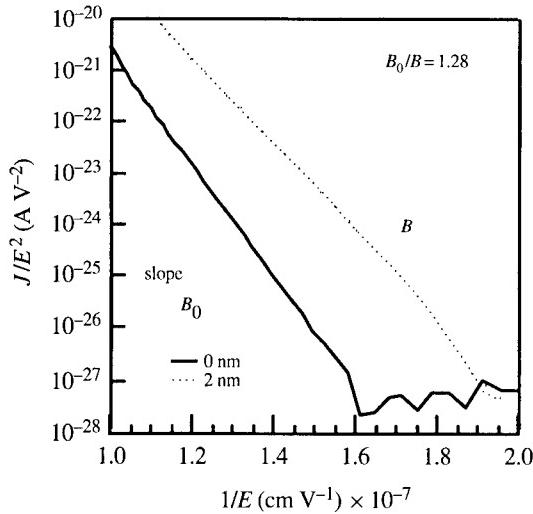
Fig. 11. F–N plot of MOS capacitors with a nm-grain poly-Si gate. The ratio B_0/B represents the local field enhancement factor.
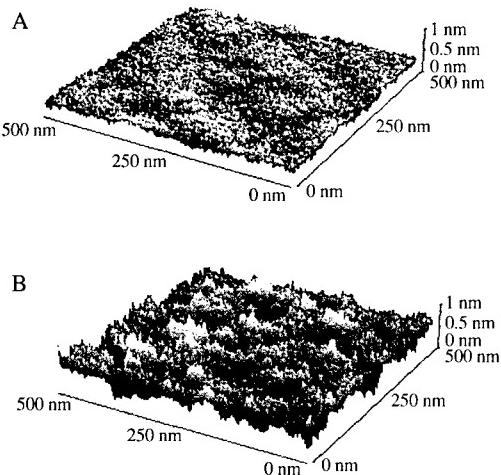
Fig. 12. AFM image of the tunnel oxide surface after poly-Si gate removal by N_2H_4 etching. A, without a nm-grain poly-Si layer. B, with a nm-grain poly-Si layer when the thickness is 4 nm.

on the thickness of the nm-grain layer (Fig. 10). Thus, this gate structure is promising in FLASH memories to increase write/erase speed without concern of retention problems [7, 13].

(iv) *Model for the oxide/gate interface structure.* To explain these phenomena, we assumed local electric field enhancement at the poly-Si/tunnel oxide interface. This assumption is supported by Fig. 11, which shows the F–N plot [10] from I–V measurement results. In this figure the ratio of slope B_0/B represents the electric field enhancement factor [16]. When the electric field enhances locally, the value B_0/B (the value is 1.28 with the nm-grain poly-Si) should be greater than 1. We also observed the tunnel oxide surface after removal of the poly-Si gate by N_2H_4 etching using AFM. As shown in Fig. 12, the tunnel oxide surface with nm-grain poly-Si have a roughened structure that may cause the local
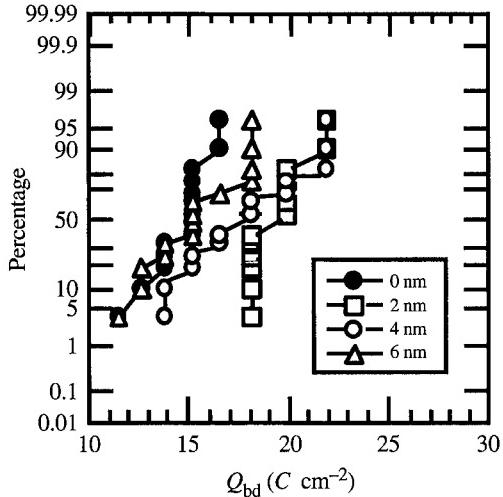
Fig. 13. Constant current TDDB measurement results for MOS capacitors with a nm-grain poly-Si double-layer gate.

electric field enhancement by the geometry effect [17]. In contrast, field enhancement does not cause in the case of positively biased.

(v) *TDDB characteristics.* In the nm-grain poly-Si/tunnel oxide structure, charge to breakdown, Q_{bd}, increased as shown Fig. 13.

F. High-k gate dielectric films

The electric properties of TiO_2 gate insulators deposited by electron-beam evaporation have been studied and a 1 nm equivalent oxide thickness can be achieved. The leakage current density of 1 nm TiO_2 is small enough, as comparable to 3.3 nm nitrided SiO_2. The k-value in TiO_2 varied widely with annealing conditions [18].

4. Conclusions

(1) We have investigated an oxide process that ensures high reliability even in the giga-scale integration generation. Wet oxidation with ultra-dry unloading combined with pre-cleaning in a closed wet cleaning system is promising for ensuring high yields even with a large total gate area. Thus, it may be possible to use ultra-thin SiO_2 film as a gate oxide in future ULSIs retaining sufficient reliability, by controlling the Si surface and moisture in the gate oxidation process.

(2) Using nm-thick a-Si film, we developed a novel nm-grain poly-Si gate structure which is suitable for reducing cell-to-cell tunnel current deviation. This structure is promising in quarter-micron high-speed FLASH memories because of the increase of tunnel current and Q_{bd} without any degradation of the tunnel oxide reliability. These phenomena are explained by local field enhancement caused by the nm-grain poly-Si interface structure.

(3) In high-k gate dielectric film (TiO_2), the k-value changed dramatically with annealing conditions. Interface structure control is a key technique to realizing high-quality high-k gate dielectric films in future devices.

References

[1] C. Hu, Proc. IEEE **81**, 682 (1993).

[2] J. C. Lee, I.-C. Chen, S. Holland, Y. Fong, and C. Hu, IEEE Trans. Electron Devices **35**, 2268 (1989).

[3] J. Yugami and M. Ohkura, in *Proceedings of the IEEE International Conference on Microelectronic Test Structures*, (1994) Vol. 7, p. 823.

[4] T. Itoga, A. Hiraoka, F. Yano, J. Yugami, and M. Ohkura, Extended Abst. of SSDM, 667 (1994).

[5] M. Hao, K. Lai, W. Chen, and J. C. Lee, IEDM Tech. Dig., 601 (1994).

[6] J. Yugami and M. Ohkura, Extended Abst. of SSDM, 270 (1995).

[7] F. J. Feigl, D. R. Young, D. J. DiMalia, S. K. Lai, and J. A. Calise, J. Appl. Phys. **52**, 5665 (1981).

[8] J. Yugami, T. Mine, S. Iijima, and A. Hiraiwa, in *Proceedings of the IEEE International Conference on Microelectronic Test Structures*, (1992) Vol. 5, p. 1.

[9] J. C. Lee *et al.*, IEEE Trans. Electron Devices **35**, 2268 (1989).

[10] S. M. Sze, *Physics of Semiconductor Devices* (Wiley Interscience, New York, 1981) 2nd edn, p. 440.

[11] G. Lucovsky *et al.*, J. Vac. Sci. Technol. **B5**, 530 (1987).

[12] D. J. Dimaria *et al.*, Appl. Phys. Lett. **61**, 2329 (1992).

[13] M. Ushiyama *et al.*, IEEE IRPS, 3331 (1991).

[14] S. Muramatsu *et al.*, IEEE Tech. Dig. IEDM, 847 (1994).

[15] T. Ishii *et al.*, Ext. Abst. of SSDM95, 201 (1995).

[16] J. Yugami *et al.*, Proc. IEEE ICMTS, 17 (1991).

[17] J. Tanaka *et al.*, Ext. Abst. of SSDM95, 240 (1995).

[18] K. Haraguchi *et al.*, Ext. Abst. of SSDM99, 376 (1999).

Tight-binding investigation of electron tunneling through ultrathin SiO$_2$ gate oxides[†]

M. STÄDELE, B. R. TUTTLE, K. HESS, L. F. REGISTER

Beckman Institute, University of Illinois, Urbana, IL 61801, U.S.A.

We investigate electron tunneling through ultrathin gate oxides using scattering theory within a tight-binding framework. We employ Si[100]/SiO$_2$/Si[100] model junctions with oxide thicknesses between 7 and 18 Å. This approach accounts for the three-dimensional microscopic structure of the model junctions and for the three-dimensional nature of the corresponding complex energy bands. The equilibrium positions of the atoms in the heterostructure are derived from first-principles density-functional calculations. We show that the present method yields qualitative and quantitative differences from conventional effective-mass theory.

© 2000 Academic Press

Key words: tunneling, thin oxides, silicon dioxide, gate dielectrics, tight-binding.

1. Introduction

Gate oxide thicknesses in metal-oxide semiconductor (MOS) transistors are approaching 2 nm. Consequently, gate tunneling currents have become a major design limitation. Accordingly, the influence of the detailed microscopic structure and composition of the oxide on the magnitude of tunneling currents becomes increasingly important. To date, most calculations of tunneling probabilities through thin SiO$_2$ layers are performed using the effective-mass (EM) or, more generally, envelope function approximations [1–3]. However, these approaches have several well-known shortcomings that will only become more problematic for ultra-thin oxides. The EM approximation includes an unrealistic band structure for electrons with energies greater than a few tenths of an eV, it relies on an essentially one-dimensional treatment of the transmission, and does not allow for ready explanation of the apparent violations of interface-parallel momentum conservation that have been observed experimentally [4]. Also, the EM approaches do not explicitly include the microscopic features of the oxide, whereas for ultrathin gate oxides these features become increasingly important.

In this paper, we present a method that addresses many of the limitations of the EM theory. We present results for electron transmission through Si[100]/SiO$_2$/Si[100] barriers, obtained from fully three-dimensional tight-binding (TB) calculations that are based on a scattering theory framework. This formalism allows one to treat the Si and SiO$_2$ band structures realistically for energies up to several eV. The atomic coordinates of the underlying microscopic model of the junction were determined by *ab initio* density-functional calculations. Implications of the present approach that are relevant for the theoretical modeling of gate currents are discussed and transmission coefficients are compared for TB and EM calculations.

[†] Devoted to the SIMD Conference in Hawaii (December 1999).

2. Computational methods

We briefly summarize the computational procedure that we have used to obtain the results for the Si[100]/SiO$_2$/Si[100] systems studied here, leaving the details for a future publication [5]. The strategy of the present work consists of two main steps.

First, realistic microscopic models of the silicon-oxide-silicon junction (with various oxide thicknesses) based on the tridymite polytype of SiO$_2$ were constructed (see also [6]). The atomic coordinates were subsequently relaxed to their equilibrium values using gradient-corrected (GGA) local-density calculations [7]. The resulting structures are characterized by sharply peaked Si–O bondlength distributions centered around 1.61 Å. Due to some residual strain, these SiO$_2$ model gate oxides are energetically slightly less favorable than α-quartz. Laterally, the basic unit cells of the models have dimensions of $\sqrt{2}a_{Si} \times \sqrt{2}a_{Si}$ ($a_{Si} = 5.43$ Å) which extends over *four* minimum unit cells of Si. To discern these models from the normal Si bulk unit cell whose dimensions are $a_{Si}/\sqrt{2} \times a_{Si}/\sqrt{2}$, we shall denote the former '2 × 2' and the latter '1 × 1'. Note that real gate oxides are known to be disordered on length scales that are larger than we can currently model. We plan to develop new models with varying lateral size and topology.

Second, using the above model as input, we obtain reflection and transmission coefficients for the whole structure by employing a TB method that is embedded in a scattering theory framework [8–10]. We solve the Schrödinger equation for the whole junction at a fixed energy E and in-plane momentum $k_{\parallel}^{2\times2}$ (here and in the following, the superscript indicates that the in-plane momentum belongs to the 2 × 2 Brillouin zone which is four times smaller than the standard Si 1 × 1 zone). The key idea of this approach is essentially detailed in [8]; we employ several refinements which allow one to deal with noninvertable transfer matrices [9] and to include the boundary conditions in an extremely straightforward manner [10].

The present approach guarantees that the proper boundary conditions for the scattering problem are satisfied, i.e. an incoming Bloch state in the Si bulk region with momentum $k_{\perp}^{in}$ normal to the interface is scattered into a set of outgoing states $\{k_{\perp,j}^{out}\}$ that propagate and decay in the same (transmission) or the opposite (reflection) direction. Transmission coefficients are obtained as

$$T(E, k_{\parallel}^{2\times2}) = \sum_i \sum_j |t(k_{\perp,i}^{in} \to k_{\perp,j}^{out})|^2 \frac{v_{\perp,j}^{out}}{v_{\perp,i}^{in}} \tag{1}$$

where $v_{\perp,i}^{in}$ and $v_{\perp,j}^{out}$ denote the components of the group velocities of the respective states normal to the interfaces; the transmission amplitude $t(k_{\perp,i}^{in} \to k_{\perp,j}^{out})$ is obtainable by solving a linear matrix equation [8, 10]. We are able to distinguish incident and transmitted Si states of identical $k_{\parallel}^{2\times2}$ that differ in $k_{\parallel}^{1\times1}$.

An sp^3 TB basis with second-nearest-neighbor interactions for both silicon [11] and the oxide [5] was used. The Si TB conduction band structures agree fairly well with experiment for energies up to 3 eV. The oxide parameters were chosen to yield a bandgap of 8.9 eV and to reproduce the GGA effective mass (0.42 m_0) of the lowest SiO$_2$ conduction band in the [100] direction.

3. Tight binding versus effective-mass theory

It is evident that the EM picture is unphysical at high (≥ 0.5 eV) incident electron energies since the concept of a single parabolic valley in the Si band structure is invalid. But even at lower energies, the full band-structure picture of electron tunneling in the present TB scheme has some important features that are qualitatively different from the standard EM approach.

The EM viewpoint is illustrated schematically in Fig. 1A which depicts the projections of the lowest conduction bands of Si and SiO$_2$ on the [001] or [010] axis. Note that the lowest conduction bands of Si consist of six degenerate ellipsoids that lie along the [100] and the corresponding symmetry-equivalent directions, $0.8 \times \frac{2\pi}{a_{Si}}$ away from $k_{\parallel} = 0$; SiO$_2$ is normally assumed to have a parabolic conduction band centered around
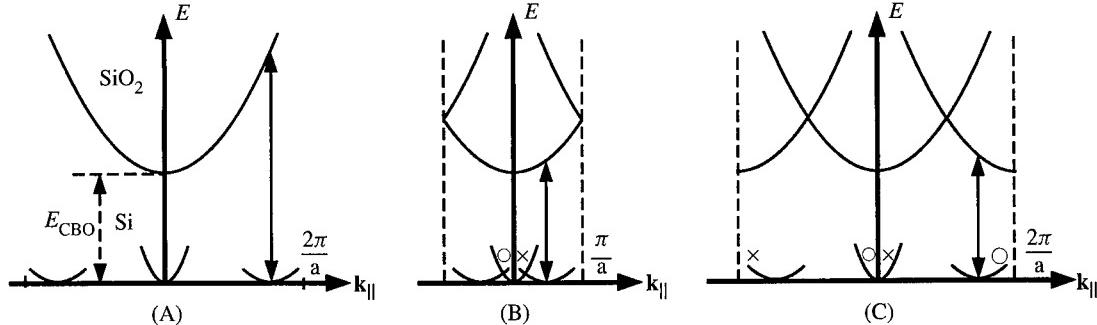
Fig. 1. Schematic views of the lowest conduction bands of Si and SiO_2, projected on an axis through $k_\parallel = 0$ that points in the [010] (or, equivalently, in the [001]) direction. The SiO_2 bands are shifted relative to the Si bands by the conduction band offset E_{CBO}. A, Effective-mass picture B, tight-binding picture for the 2 × 2 structure studied here: the bulk Si bands are folded in the 2 × 2 Brillouin zone; the cross and circle indicate regions where electrons from different valleys couple. C, is the same as B, but outfolded back in a 1 × 1 cell.

$k_\parallel = 0$. Consequently, conservation of in-plane momentum $k_\parallel$ in the EM picture requires that electrons in the [010], [0$\bar{1}$0], [001], and [00$\bar{1}$] minima with very-large in-plane momenta $k_\parallel \approx 0.8 \times \frac{2\pi}{a_{Si}}$ (called 'off-axis electrons' hereafter) see a much higher effective barrier for tunneling than the [100] and [$\bar{1}$00] electrons with $k_\parallel \approx 0$ (termed 'on-axis electrons'). Analogously, the EM picture implies that the barrier to transport from the indirect semiconductor Si into SiO_2 should similarly be a strong function of the orientation of the Si to the interface.

However, little if any orientation dependence is found experimentally in the barrier to tunneling [4, 12–14], suggesting a strong $k_\parallel$-breaking interaction across the interface. To date, this apparent lack of $k_\parallel$ conservation has been ascribed to disorder in the oxide, interface roughness [4, 15], and phonon scattering [15]. Accordingly, it also has been argued that to an extent there is short-range order in the oxide, the EM picture with parallel momentum conservation remains valid [2]. However, the present model is highly ordered, yet still allows for $k_\parallel$-breaking. Consider the 2 × 2 models studied here. Their projected Brillouin zones are four times as small as that of bulk Si (the boundaries in [010] direction lie at $k_\parallel^{2 \times 2} = \frac{\pi}{a_{Si}}$, see Fig. 1B) and thus, the four off-axis ellipsoids are reduced into this smaller zone by a reciprocal lattice vector $G_\parallel$ of the 2 × 2 zone. This band-folding has at least two important implications: (i) electrons in the off-axis ellipsoids see a much different barrier than the barrier that follows from the EM model (compare Figs 1A and B) and (ii) Si bulk states with different $k_\parallel$ (relative to the bulk 1 × 1 Brillouin zone) couple at the Si–SiO_2 interface because of band folding, producing an inherent violation of parallel momentum conservation with respect to the Si Bloch states. Of course, the true oxide is also disordered, so the band-folding effects could be far more pronounced than in the ordered model of this work. These band-folding effects could have significant implications for modeling tunneling through and transport over the Si/SiO_2 interface in MOS devices where both $k_\parallel$-conserving [1–3] and $k_\parallel$-breaking [16, 17] models are still commonly used depending on the *assumed* relative importance of the $k_\parallel$-breaking transport channel.

4. Quantitative results for transmission coefficients

In this section, we compare TB and EM transmission coefficients quantitatively. For the EM calculations, the effective masses obtained from the TB models have been used.

Figure 2A shows the transmission coefficients obtained for an on-axis state with $k_\parallel = 0$ and an energy $E = 0.02$ eV relative to the conduction band minimum in Si for various oxide thicknesses and bias voltages. The EM and TB results differ by one order of magnitude or less for bias voltages ranging from 0–2.5 eV; both
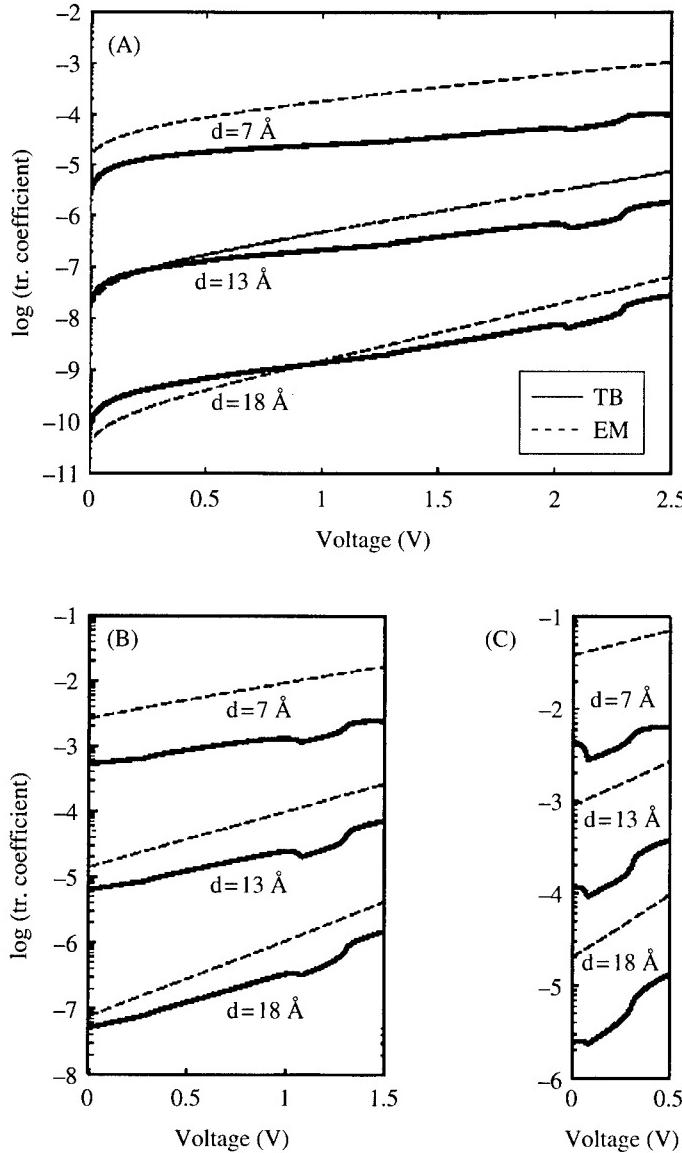
Fig. 2. Transmission coefficients $T(E, k_\parallel = 0)$ for incident electron energies E of: A, 0.02 eV, B, 1.0 eV, and C, 2.0 eV as a function of applied voltage, obtained within the present tight-binding (solid lines) and the effective-mass (dashed lines) methods for various oxide thicknesses d. Note that the B and C depict the sums of the TB transmission coefficients of all six and ten incoming states, respectively.

models correctly obtain the roll-off region for small voltages < 0.2 eV. The EM curves are generally flatter, indicating the presence of larger complex wavevectors in the SiO_2 gap. In addition, the TB results show band structure effects (slight 'kinks' at energies where the states around $k_\parallel = 0$ into which the incoming electrons are scattered change their number and/or character). The EM transmission coefficients tend to be larger than the TB ones for higher voltages and thinner oxide when there is stronger coupling across the oxide barrier. This trend persists as one goes to higher electron energies, as can be seen from Figs 2B and C which refer to

$E = 1.0$ and 2.0 eV, respectively. Note that for these two values of energy, we observe six and ten incoming states with $k_\parallel^{2\times2} = 0$, respectively, whereas the EM curve involves only one state.

We have also calculated the transmission coefficients for electrons incident in the off-axis ellipsoids of Si. Depending on the oxide thickness, the resulting transmission probability is about one to three orders of magnitude lower than for electrons in the on-axis ellipsoids, indicating a strong selectivity for states tunneling from on-axis valleys for this highly ordered system. However, we do see coupling between electrons in on- and off-axis valleys with the same $k_\parallel^{2\times2}$, consistent with the discussion of Section 3. For more folding due to different Si orientations or less order in the model oxide, this coupling could be far more pronounced.

5. Summary

We have studied electron tunneling through ultrathin SiO_2 slabs sandwiched between two Si [001] reservoirs within a tight-binding framework that fully accounts for the three-dimensional nature of the tunneling problem. The tight-binding picture involves important qualitative, as well as quantitative, differences from the standard effective-mass scheme. The present work is an initial step toward a microscopic understanding of tunneling in ultrathin oxides.

References

[1] F. Rana, S. Tiwari, and D. A. Buchanan, Appl. Phys. Lett. **69**, 1104 (1996).

[2] M. V. Fischetti, S. E. Laux, and E. Crabbé, J. Appl. Phys. **78**, 1058 (1995).

[3] C. Bowen, C. L. Fernando, G. Klimeck, A. Chatterjee, D. Blanks, R. Lake, J. Hu, J. Davis, M. Kulkarni, S. Hattangady, and I.-C. Chen, *IEDM Technical Digest (Cat. No. 97CH36103)*, New York, (1997) p. 869.

[4] Z. A. Weinberg, J. Appl. Phys. **53**, 5052 (1982).

[5] M. Städele, B. R. Tuttle, and K. Hess, unpublished.

[6] A. Pasquarello, M. S. Hybertsen, and R. Car, Appl. Phys. Lett. **68**, 625 (1996).

[7] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer, Berlin, 1990).

[8] D. Z.-Y. Ting, E. T. Yu, and T. C. McGill, Phys. Rev. **B45**, 3583 (1992).

[9] T. B. Boykin, Phys. Rev. **B54**, 7670 (1996).

[10] C. Strahberger, private communication.

[11] G. Grosso and C. Piermarocchi, Phys. Rev. **B51**, 16772 (1995).

[12] Z. A. Weinberg, Solid State Electron. **20**, 11 (1977).

[13] Z. A. Weinberg and A. Hartstein, J. Appl. Phys. **54**, 2517 (1983).

[14] G. Krieger and R. M. Swanson, Appl. Phys. Lett. **39**, 818 (1981).

[15] U. Kunze and G. Lautz, Surf. Sci. **113**, 55 (1982).

[16] L. F. Register, E. Rosenbaum, and K. Yang, Appl. Phys. Lett. **74**, 457 (1999).

[17] S.-H. Lo, D. A. Buchanan, Y. Taur, and W. Wang, IEEE Elec. Dev. Lett. **18**, 209 (1999).

Effect of single-electron interface trapping in decanano MOSFETs: A 3D atomistic simulation study

A. ASENOV[†], R. BALASUBRAMANIAM, A. R. BROWN, J. H. DAVIES

Device Modelling Group, Department of Electronics and Electrical Engineering, The University of Glasgow, Glasgow G12 8LT, U.K.

We study the effect of trapping/detrapping of a single-electron in interface states in the channel of n-type MOSFETs with decanano dimensions using 3D atomistic simulation techniques. In order to highlight the basic dependencies, the simulations are carried out initially assuming continuous doping charge, and discrete localized charge only for the trapped electron. The dependence of the random telegraph signal (RTS) amplitudes on the device dimensions and on the position of the trapped charge in the channel are studied in detail. Later, in full-scale, atomistic simulations assuming discrete charge for both randomly placed dopants and the trapped electron, we highlight the importance of current percolation and of traps with strategic position where the trapped electron blocks a dominant current path.

© 2000 Academic Press

Key words: MOSFET, trapping, interface, 3D atomistic simulation.

1. Introduction

With the scaling of MOSFETs to decanano dimensions towards the end of the silicon roadmap [1], current fluctuations in the form of random telegraph signals (RTS), caused by trapping/detrapping of single carriers in interface states in the channel region [2], are becoming increasingly important. Recently, RTS with amplitudes larger than 60% have been observed at room temperature in very narrow channel MOSFETs [3]. Current fluctuations on such a scale will become a serious issue, not only for the next generation of analogue circuits, but also for mixed-mode [4] and digital applications.

It is widely accepted that RTS originate from localized modulation of the electrostatic potential induced by the trapped charge, and corresponding local reduction in carrier density and/or mobility [2–4]. There is, however, a lack of both analytical models and simulation studies explaining the wide range of RTS amplitudes in otherwise identical devices [5]. For example, the existing analytical models [6] and quasi-2D simulations [7] do not include an influence of the trap position along the channel on the RTS amplitude. It has also been suggested that strategically located traps influence the magnitude and the spreading of the RTS amplitudes due to surface potential fluctuations, and the corresponding current percolation [2, 8]. This has been confirmed [9] by using 2D numerical simulations in the plane of the MOSFET channel and assuming that the source of the potential fluctuations are a random fixed interface and oxide charges. However, in the next

---

[†]E-mail: A.Asenov@elec.gla.ac.uk

generation MOSFETs with decanano dimensions the surface potential fluctuations will be dominated by the random distribution of discrete dopants in the channel depletion region [10] and the complete investigation of the RTS amplitudes in realistic devices requires full-scale 3D numerical simulations.

In this paper we use an efficient 3D atomistic simulation technique [11] to study the RTS amplitudes associated with trapping/detrapping of a single electron in an interface state in the channel of decanano MOSFETs. For the first time an enhancement in the RTS amplitudes associated with random dopant-induced surface potential fluctuations, and the corresponding current percolation, is captured in our simulations.

2. Simulation approach

We investigate the change in the drain current associated with trapping of an electron in a single-acceptor-type interface state in the presence of random discrete dopants in n-channel MOSFETs. The atomistic simulation technique used in this study is described in detail elsewhere [11]. The simulations are restricted to low drain voltage and are based on a single 3D solution of the nonlinear Poisson equation where the concentrations of electrons and holes follow Boltzmann statistics. The current is calculated from the channel resistance obtained by solving a simplified current continuity equation in a drift approximation only. A uniform grid with typical grid spacing $h = 1$ nm is used in the discretization of the Poisson equation in order to resolve the effects associated with individual dopants and single trapped electrons. The discrete dopant or trapped electron charge is introduced in the solution of the Poisson equation by assigning a doping concentration of $1/h^3$ to the nearest grid node.

At this stage we study the effects associated only with the local change in the carrier density assuming that the mobility is constant. All results presented are for decanano n-channel MOSFETs with a square geometry. The investigated devices have oxide thickness $t_{ox} = 3$ nm, uniform doping concentration in the channel region $N_A = 5 \times 10^{18}$ cm^{-3} and junction depth $x_j = 7$ nm with 5 nm lateral sub-diffusion. The above choice of MOSFET parameters allows for a direct comparison with previously published atomistic simulation results [11].

3. Continuous doping

To highlight the major trends and dependencies, we initially consider a continuous charge representing the doping concentration in the channel and in the source/drain regions of the MOSFETs under investigation. A single, localized electron charge q is introduced at the Si/SiO$_2$ interface to represent an electron trapped in an acceptor-type interface state.

The relative change in the drain current (relative RTS amplitude) associated with the trapping of a single electron in the middle of the channel of a 100×100 nm MOSFET is plotted in Fig. 1 as a function of the drain current in a double-logarithmic scale. At room temperature, the relative RTS amplitude shows a constant plateau value for low drain current (weak inversion) and roll-off which is proportional to I_D^{-1} in strong inversion. This is in agreement with the experimental observations in [6] and [8]. The roll-off region is related to the screening of the influence of the trapped charge on the potential distribution by electrons in the inversion layer, a process which is naturally included in the solution of the nonlinear Poisson equation. It is important to point out that the electrostatics associated with the screened potential of the trapped charge is sufficient to explain the experimentally observed behaviour of the RTS amplitude even if a constant mobility is used in the simulations.

The dependence of the relative RTS amplitude on the gate voltage for a set of decanano devices covering the whole range of device dimensions until the end of the silicon roadmap is presented in Fig. 2. A dashed line, corresponding to the threshold voltage in each device, separates the weak and strong inversion regions. The use of a linear scale reveals that in the weak inversion region the RTS amplitude passes through a
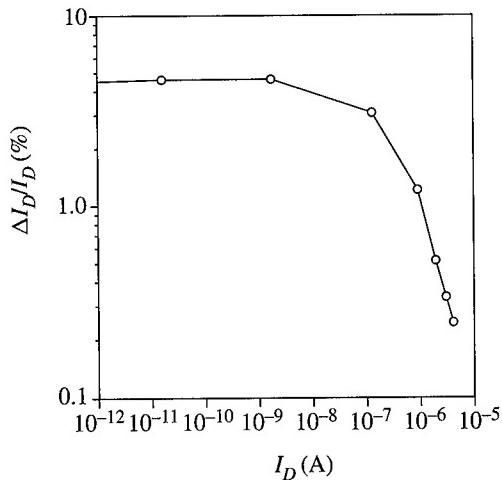
Fig. 1. Relative RTS amplitude associated with the trapping of a single electron in the middle of the channel of a 100×100 nm MOSFET as a function of the drain current. $V_D = 10$ mV.
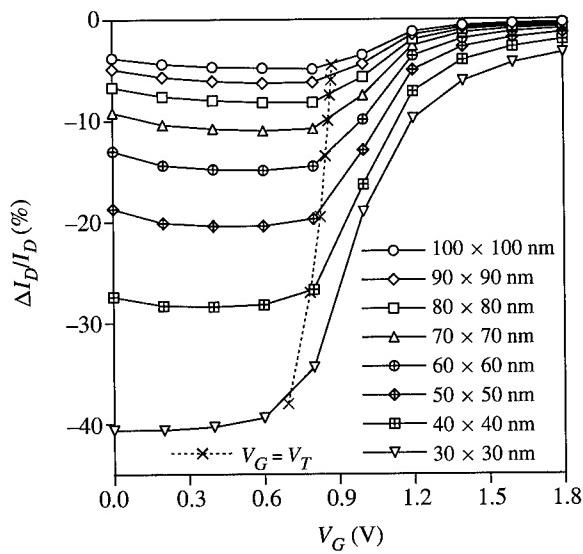
Fig. 2. Dependence of the relative RTS amplitude on the gate voltage for a set of square decanano MOSFETs. $V_D = 10$ mV.

very broad, weak maximum before rolling-off in strong inversion. The 100×100 nm MOSFET, due two generations from now, is the starting point of our investigations. The RTS amplitude for this device already reaches 5% in weak inversion but falls to a fraction of this percentage at large gate voltage. For the whole range of investigated square devices, the relative RTS amplitude remains inversely proportional to the channel area. For the 30×30 nm MOSFET, which marks the end of the roadmap, the RTS amplitude increases to more than 40% in weak inversion and remains above 5% for the whole range of gate voltage relevant to this generation of devices.

In the next simulation experiment we investigate how the position of the trapped electron affects the RTS
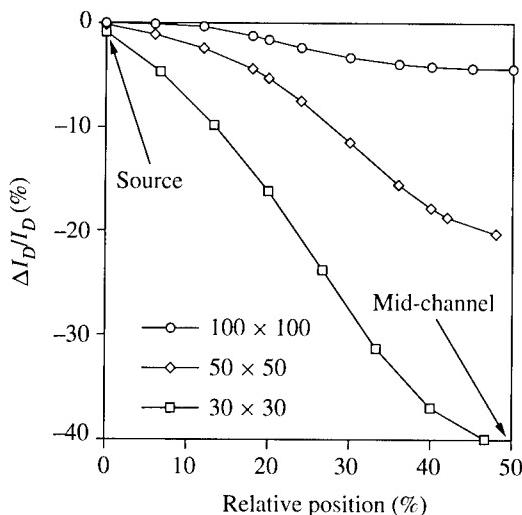
Fig. 3. Relative RTS amplitude as a function of the relative position of the trapped electron moving from the middle of the channel to the source.

amplitude by moving the trapped charge from the centre of the channel towards the source along a line perpendicular to the source p–n junction. Figure 3 illustrates the dependence of the RTS amplitude on the relative position from the middle of the channel for three MOSFETs with different dimensions. The relative RTS amplitude for a MOSFET with a particular geometry depends strongly on the position of the trapped electron. The effect of the trapped electron is negligible if the charge is close to the source/drain and reaches a maximum when the trapped electron is in the middle of the channel. The dynamic range of the position-dependent relative RTS amplitude is approximately two orders of magnitude and offers an explanation for the wide range of RTS amplitudes observed experimentally in identical devices [12]. In the longest 100×100 nm MOSFET in Fig. 3 the position dependence has a plateau near the middle of the channel and falls down in the source region where the built-in junction potential controls and raises the surface potential. Both the higher carrier concentration and conductivity of the region near the source/drain, and the corresponding screening, reduce the effect of the trapped charge on the overall current. In shorter devices the relative portion of the channel controlled by the source/drain increases and the plateau becomes narrower leading to a well-defined maximum of the RTS amplitudes in the middle of the channel for the shortest 30×30 nm MOSFET in Fig. 3.

4. Discrete random dopants

In properly scaled decanano MOSFETs the random, discrete dopants are the major factors introducing significant surface potential fluctuations and variation in the device parameters even at room temperature [11]. In a 50×50 nm device there are, on average, 170 dopants in the channel depletion region. Their actual number follows a Poisson distribution and their positions are random. The fixed oxide charge has a negligible effect in respect of the potential fluctuations compared with the random dopants. In well-controlled technology the surface density of the fixed charge is below 10^{11} cm^{-2} which is equivalent to approximately two discrete charges in every 50×50 nm MOSFET, compared with approximately 10 random dopants at the interface.

The current in the presence of strong potential fluctuations percolates through the 'valleys' in the potential landscape. The trapping and detrapping of a single charge in a strategically placed interface state can block
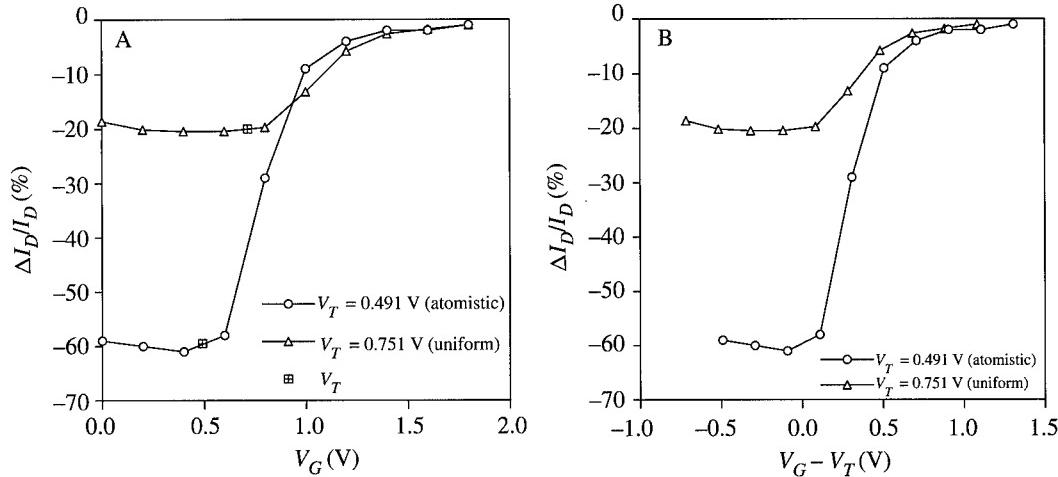
Fig. 4. A, Gate voltage dependence of the largest relative RTS amplitude in a 50 × 50 nm MOSFET with random doping and the highest relative RTS amplitude in a device with continuous doping; B, the same as A but with the two curves aligned in respect of the threshold voltage.

a dominant current path and can introduce RTS with a much larger amplitude than the maximum amplitude deduced from continuous doping simulations. In this section we present results from a full-scale atomistic simulation where not only the trapped electron charge but also the acceptors in the channel are represented as random discrete charges. The discrete random dopants are introduced only in the channel region between the source and drain with continuous doping assumed in the source and drain. The average number of dopants in these regions are calculated by multiplying the volume of the random dopant region by the corresponding doping concentrations. The actual number of dopants in each atomistic region is chosen randomly from a Poisson distribution with a mean equal to the corresponding average dopant numbers.

Unlike the simulations assuming continuous doping, the largest RTS amplitudes in the random dopant case are not in the middle of the channel but in the region with the deepest valley in the potential landscape corresponding to the highest density of percolating current.

The gate voltage dependence of the highest relative RTS amplitude in a 50 × 50 nm MOSFET with random doping and the highest relative RTS amplitude in a device with continuous doping are compared in Fig. 4A. At low gate voltages the RTS amplitude in the device with discrete dopants is three times higher and reaches 60%. The two curves, however, cross over at approximately 0.9 V. This is related to the threshold voltage lowering associated with the discreteness of the dopants captured by the atomistic simulations [11]. If the curve for the device with discrete doping is displaced by the difference in the threshold voltage to align the weak and strong inversion regions in the both devices as shown in Fig. 4B it becomes clear that in both regions the relative RTS amplitude in the case of discrete dopants remains consistently higher. However, at higher gate voltages, corresponding to strong inversion, the difference becomes smaller due to the screening of the potential fluctuations by the inversion layer charge and the higher uniformity of the corresponding current in the channel.

5. Conclusions

Trapping and detrapping of a *single* charge will have a dramatic effect on the operation of decanano MOSFETs near the end of the silicon roadmap. The relative RTS amplitudes will reach several tens of per cent when below threshold, reducing to a few per cent when above threshold, in devices with a square geometry.

This estimate is based only on the local modulation of the conducting charge in the channel and does not include the effect of the trapped charge on the channel mobility. The potential fluctuations associated with the random distribution of discrete dopants in the channel region, and the corresponding percolating current, will substantially enhance the RTS amplitudes due to the profound effect of traps strategically located in regions of high local current density.

References

[1] The National Technology Road-map for Semiconductors, Semiconductor Industry Association, San Jose, CA, 1997 Revision.

[2] K. S. Rals, W. L. Scokpol, L. D. Jakel, R. E. Howard, L. A. Fetter, R. W. Epworth, and D. M. Tennant, Phys. Rev. Lett. **63**, 228 (1984).

[3] Y. Shi, H. M. Bu, X. L. Yuan, and Y. D. Zheng, Random telegraph signals in very narrow channel MOS-FET, in *Workshop Abstracts, Silicon Nanoelectronics Workshop, Kyoto*, (1999) p. 28.

[4] S. T. Martin, G. P. Li, E. Worley, and J. White, IEEE Electron Devices Lett. **18**, 444 (1997).

[5] M.-H. Tsai and T.-P. Ma, IEEE Trans. Electron Devices **41**, 2061 (1994).

[6] E. Simoen, B. Dierick, C. L. Claeys, and G. J. Declerck, IEEE Trans. Elect. Dev. **39**, 422 (1992).

[7] A. Godoy, F. Gamiz, A. Palma, J. A. Jimenez-Tejada, and J. Banqueri, J. Appl. Phys. **82**, 4621 (1997).

[8] M. J. Uren, D. J. Day, and M. J. Kirton, Appl. Phys. Lett. **47**, 1195 (1985).

[9] H. H. Mueller and M. Schulz, J. Appl. Phys. **83**, 1734 (1988).

[10] A. Asenov, Nanotechnology **10**, 153 (1999).

[11] A. Asenov, IEEE Trans. Electron Devices **45**, 2505 (1998).

[12] Z. Shi, J.-P. Mieville, and M. Dutoit, IEEE Trans. Electron Devices **41**, 1161 (1994).

Full-band CA/Monte Carlo modeling of ultrasmall FETs

S. J. Wigger, S. M. Goodnick

Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-5706, U.S.A.

M. Saraniti

Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL 60616-3793, U.S.A.

The modeling of ultra-small MOSFETs is presented using a newly developed full-band, hybrid ensemble Monte Carlo (EMC)-cellular automata (CA) device simulator. In this hybrid approach charge transport is simulated using the CA in regions of momentum space where most scattering events occur and the EMC elsewhere, thus optimizing the trade-off between the fast, but memory-consuming CA method and the slower EMC method. In order to efficiently model 3D ultra-small FETs, the hybrid algorithm is coupled self-consistently with a 2D and 3D multi-grid Poisson solver.

© 2000 Academic Press

Key words: Monte Carlo, device simulation, transport.

1. Introduction

Particle-based methods, such as the ensemble Monte Carlo (EMC) method and more recently the cellular automata (CA) method, have demonstrated success in simulating carrier transport in semiconductor devices [1, 2]. Unfortunately, these techniques can be limited in their application due to the computational overhead in simulating devices with characteristic dimensions on the sub-micron range using full-band representation of the particle dynamics. In order to accurately model the electrical behavior of these ultra-small devices, the electronic structure and phonon spectra need to be represented with a full-band energy–momentum relation and full 3D carrier motion needs to be included. The implementation of this more accurate physical description increases the computational burden of the EMC method, making it slow and inefficient. The CA method was developed to reduce this computational burden [2, 3], and although faster, it can often require an unrealistic amount of memory. In this work, a hybrid CA/MC method is presented, which implements the CA method in regions of momentum space where most of the scattering events occur, and the MC method elsewhere, thus optimizing the relationship between memory and speed. This hybrid dynamics simulator is then coupled self-consistently with 2D and 3D Poisson solvers using a multi-grid algorithm, and a small-channel MOSFET is simulated to demonstrate the potential of this new method.

2. Hybrid CA/MC simulator

The algorithm for the CA method in k-space differs from the EMC method only in how it calculates the final state after a scattering event. The EMC approach stores the scattering probability integrated over the
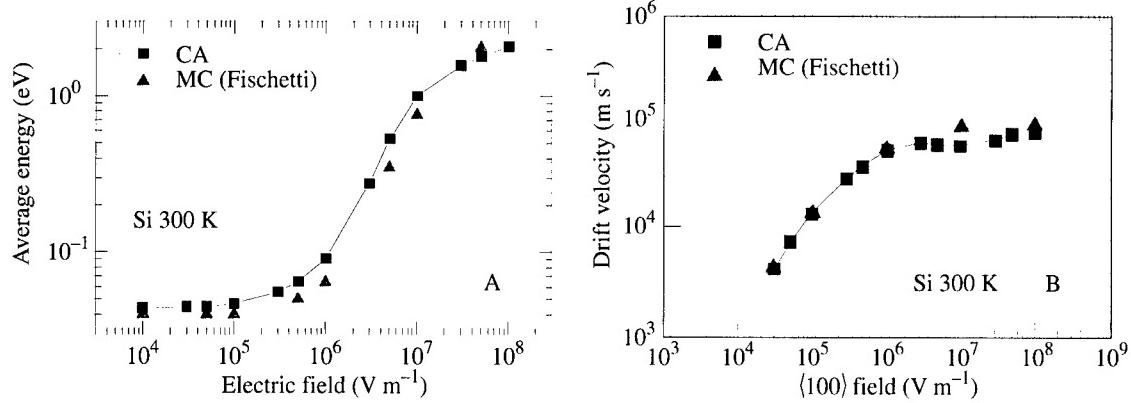
Fig. 1. Comparison of the CA method with EMC for Si at $T = 300$ K for the A, the drift velocity and B, the average energy vs. field.
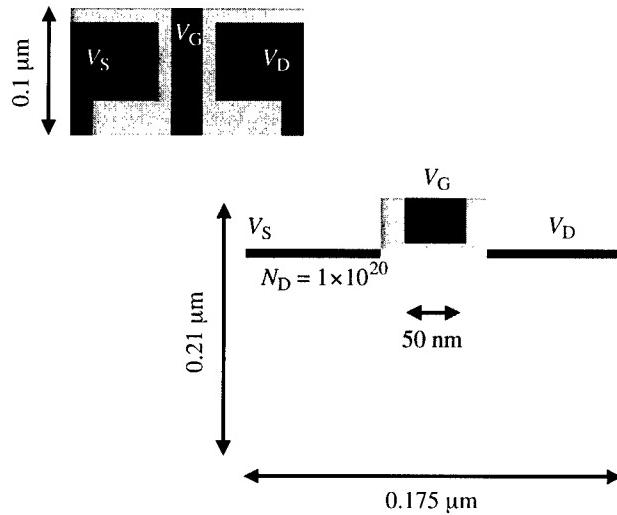
Fig. 2. Three-dimensional schematic layout for a 50 nm n-channel MOSFET.

whole momentum space. The final state after a scattering event is then obtained by inverting the energy–momentum dispersion relation, which is tabulated when the full-band representation is implemented. Within the CA approach, the probabilities of scattering from an initial momentum state to all possible final states satisfying the appropriate conservation laws are computed and tabulated. This greatly reduces the computational demand, since the final state can be chosen with a single random number. The drawback of this method is that, since all the final states need to be tabulated, an appreciable amount of memory is required to store all of these transition tables. Some of the memory requirements are alleviated using a nonuniform grid in the first Brillouin zone. These memory requirements can be further reduced by combining the CA and MC methods. In this hybrid approach, the MC scattering tables are implemented in regions where the carrier population tends to be small, and/or scattering probabilities are low. This framework allows for the use of the fast, but memory-consuming CA in the most active areas of momentum space, while minimizing the amount of required memory by using the MC scattering selection and final-state computation elsewhere.
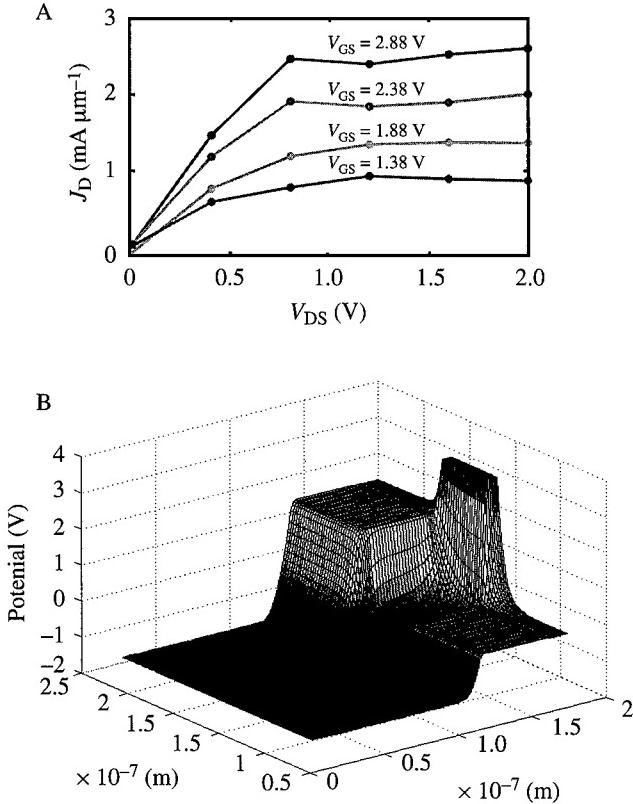
Fig. 3. A, $I-V$ characteristics for MOSFET structure using the 2D Poisson solver. B, Potential profile using the 3D hybrid CA/multigrid algorithm for $V_g = 2.88$ V and $V_d = 2.0$ V.

When used in this hybrid mode, the simulator can adaptively update the transition/scattering tables to account for changes in the location of high scattering regions.

In the present work, acoustic phonons, nonpolar optical phonons and ionized impurity scattering are included. The models used are essentially identical to those used in Fischetti and Laux [4], in order to compare the results for the full-band CA (see Fig. 1). Impact ionization is also included with a separate algorithm due to the multi-particle nature of this mechanism.

The energy-dispersion relation and the phonon spectra are calculated across the entire Brillouin zone using an inhomogeneous grid. Due to symmetry, only values in the irreducible wedge of the first Brillouin zone actually need to be determined and stored. The band structure is calculated using the nonlocal empirical pseudopotential method (EPM) [5] including spin–orbit interaction, and the phonon dispersion is calculated using an empirical shell model [6]. Both electrons and holes are simulated.

Figure 1 shows the comparison of the velocity-field and average energy vs. field of the hybrid simulator compared with the results of Fischetti and Laux using the full-band EMC simulator DAMACLES. As can be observed, the comparison is good overall, validating the hybrid CA approach.

3. Poisson solver

Both 2D and 3D multi-grid Poisson solvers are self-consistently coupled with the hybrid CA/MC algorithm. Since motion is now treated in three dimensions, the amount of time necessary for the repeated solution of Poisson's equation can be very large and the implementation of an efficient iterative solver is essential. The multi-grid method is one of the most efficient iterative techniques available to solve large, sparse systems of equations [7]. Most iterative solvers work by reducing the high-frequency components of the error, leaving the low-frequency components unchanged. The result is a convergence rate that is initially quick, but slows down after the high-frequency components have been reduced. The multi-grid method, on the other hand, utilizes several grids with varying mesh sizes, in order to smooth the different Fourier components of the error simultaneously, resulting in a dramatic speed up.

The velocities of both electrons and holes are determined from the solution of Poisson's equation, which is written in a generic way, so that different device structures can be simulated. Input parameters include inhomogeneous dielectric constants, surface charges, inhomogeneous mesh spacing, irregular boundary conditions, and complex geometries.

4. Simulation results and discussion

Figure 2 shows the top schematic view of a 50 nm gate length device structure corresponding to a process currently being developed at Arizona State. The doping in the channel and substrate is 1×10^{19} cm^{-3} p-type while the source and drain contacts are 1×10^{20} cm^{-3} n-type. The gate oxide thickness is 3 nm.

To date, full I–V characteristics have only been obtained using the 2D solver due to the still present computational burden of simulating both electrons and holes in this MOSFET structure using the full-band CA/MC algorithm discussed here. The I–V characteristics for four gate bias are shown in Fig. 3A, which appear reasonable for this doping density and gate length. The speed of the present combined algorithm is found to be substantially greater than a standard EMC/2D-Poisson based on the same multi-grid algorithm, although neither the standard EMC nor the hybrid CA/MC has been optimized for maximum performance, so more comparison is necessary. The results of the 3D simulator for the potential distribution inside the device for a fixed gate and drain bias in the active region is shown in Fig. 3B. Again, the expected potential profile is found, demonstrating the feasibility of the approach.

References

[1] C. Jacaboni and L. Reggiani, Rev. Mod. Phy. **55**, 645 (1983).

[2] K. Kometer, G. Zandler, and P. Vogl, Phys. Rev. **B46**, 1382 (1992).

[3] M. Saraniti, S. J. Wigger, and S. M. Goodnick, Full-Band cellular automata for modeling transport in submicron devices, in *Proceedings of the Second Int. Conf. on Modeling and Simulation of Microsystems*, (1999) pp. 380–383.

[4] M. V. Fischetti and S. E. Laux, Phys. Rev. **B38**, 9721 (1988).

[5] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. **B14**, 556 (1976).

[6] K. Kunc and O. H. Neilson, Comput. Phys. Commun. **17**, 413 (1979).

[7] W. Hackbush, *Multi-grid Methods and Applications* (Springer-Verlag, Berlin, 1985).

Interface-charged impurity scattering in semiconductor MOSFETs and MODFETs: temperature-dependent resistivity and 2D 'metallic' behavior

S. Das Sarma, E. H. Hwang, Igor Žutić

Department of Physics, University of Maryland, College Park, Maryland, 20742-4111, U.S.A.

We present the results on the anomalous 2D transport behavior by employing Drude–Boltzmann transport theory and taking into account the realistic charge impurity scattering effects. Our results show quantitative agreement with the existing experimental data in several different systems and address the origin of the strong and nonmonotonic temperature-dependent resistivity.

© 2000 Academic Press

Key words: metal–insulator semiconductor structures, metal–insulator transition.

1. Introduction

A large number of recent experimental publications on low-temperature transport measurements in low-density high-mobility two-dimensional (2D) electron systems in Si MOSFETs [1], GaAs MODFETs [2], and SiGe hetrostructures [3] report an anomalously strong temperature-dependent resistivity in the narrow regime of 0.1–5 K. In contrast to the usual Bloch–Grüneisen theory of essentially a temperature-independent low-temperature resistivity, the measured resistivity changes by as much as a factor of 10 for a 1–2 K increase in temperature. This observed anomaly has led to a great deal of theoretical activity [4, 5] involving claims of an exotic metal or even a superconducting system at the interface producing the strong temperature-dependent resistivity, which has no known analog in ordinary 3D metallic behavior. Much more interest has focused around the possibility of a 2D metal–insulator quantum phase transition being responsible for the observed strong temperature-dependent resistivity since theoretically a 2D electron system at $T = 0$ has so far been thought to be (at least in the absence of electron interaction effects) an insulator [6].

In this paper we provide a theoretical explanation for the temperature-dependent resistivity of the 2D systems in the 'metallic' phase ($n_s \geq n_c$, where n_s is the 2D density and n_c the critical density which separates 'metallic' and 'insulating' behavior) in the absence of magnetic field [7] by using the Drude–Boltzmann transport theory with RPA screening and the Dingle temperature approximation to incorporate collisional broadening effects on screening [8]. In our approach we leave out quantum corrections, including localization effects, and neglect the inelastic electron–electron interaction, which may well be significant in the low-density 2D systems of experimental relevance. Our calculated resistivity agrees quantitatively with the existing experimental data [1–3] on the temperature-dependent low-density resistivity of 2D electron systems. We find that the strong temperature dependence arises from a combination of two effects: the strong temperature dependence of finite wavevector screening in 2D systems and a sharp quantum classical crossover due to the low Fermi temperature in the relevant 2D systems.
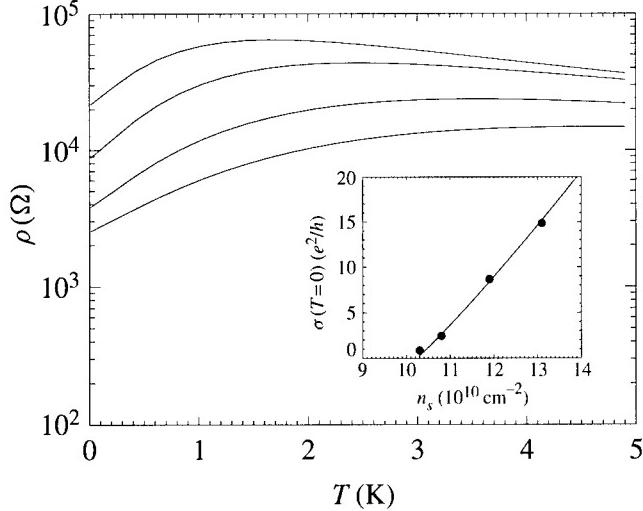
Fig. 1. The calculated resistivities for various electron densities, $n_s = 1.03, 1.08, 1.19, 1.31 \times 10^{11}$ cm^{-2} (top to bottom) as a function of T for the Si-12 sample of Ref. [1], using the critical density $n_c = 10^{11}$ cm^{-2}. In the inset we show the analytic zero-temperature conductivity as a function of density n_s. Points represent extrapolated $\sigma(T \to 0)$ from Si-12 sample of Ref. [1].

2. Theory

We use the finite temperature Drude–Boltzmann theory to calculate the ohmic resistivity of the inversion layer electrons, taking into account only long-range scattering by the static charged impurity centers with the screened electron-impurity Coulomb interaction. The screening effect is included within the random phase approximation (RPA) with the finite temperature static RPA dielectric (screening) function $\kappa(q, T)$ given by

$$\kappa(q, T) = 1 + \frac{2\pi e^2}{\bar{\kappa} q} F(q) \Pi(q, T),$$
(1)

where $F(q)$ is the form factor for electron–electron interactions and $\Pi(q, T)$ is the static polarization. We assume that the charged impurity centers are randomly distributed in the plane parallel to the semiconductor–insulator surface. Within the Born approximation the scattering time $\tau(\varepsilon, T)$ for our model is given by

$$\frac{1}{\tau(\varepsilon, T)} = \frac{2\pi}{\hbar} \int \frac{d^2 k'}{(2\pi)^2} \int_{-\infty}^{\infty} N_i(z) dz \left| \frac{v_q(z)}{\kappa(q, T)} \right|^2 (1 - \cos\theta) \delta(\epsilon_k - \epsilon_{k'}),$$
(2)

where $q = |k - k'|$, $N_i(z)$ is the impurity density of the charged center, $\theta \equiv \theta_{kk'}$ is the scattering angle between k and k', $\varepsilon = \epsilon_k = \hbar^2 k^2 / 2m$, $\epsilon_{k'} = \hbar^2 k'^2 / 2m$, $v_q(z)$ is the 2D electron–impurity Coulomb interaction. In calculating the Coulomb interaction and the RPA dielectric function in eqn (1) we take into account subband quantization effects in the inversion layer through the lowest subband variational wavefunction. The resistivity is given by $\rho = m/(ne^2 \langle \tau \rangle)$, where m is the carrier effective mass, n the effective free carrier density [5], and $\langle \tau \rangle$ the energy-averaged scattering time. The average is given by $\langle \tau \rangle = \int d\varepsilon \tau(\varepsilon) \varepsilon \left(-\frac{\partial f}{\partial \varepsilon} \right) / \int d\varepsilon \left(-\frac{\partial f}{\partial \varepsilon} \right) \varepsilon$, where $f(\varepsilon)$ is the Fermi distribution function, $f(\varepsilon) = \{1 + \exp[(\varepsilon - \mu)]/k_B T\}^{-1}$ with finite temperature chemical potential, $\mu = \mu(T, n)$, which is determined self-consistently.

3. Results and conclusion

It is physically instructive to first consider the asymptotic behavior of the temperature-dependent part of resistivity, $\rho(T)$. In the quantum regime at low temperature, $T \ll T_F$ with $T_F \equiv \mu(T = 0)/k_B$, the dominant behavior of $\rho(T)$ is linearly increasing with T, i.e. $\rho(T) \propto T/T_F$ arising from the temperature-dependent screening, $\kappa(q, T)$ [8]. In the high-temperature limit $(T \gg T_F)$ corresponding to the classical regime, the resistivity is decreasing with temperature, i.e. $\rho(T) \propto T_F/T$ due to the energy averaging of τ. For intermediate temperatures $(T \sim T_F)$ the system crosses over from a nondegenerate classical to a strongly screened degenerate quantum regime [5].

In Fig. 1 we present our numerically calculated resistivity $\rho(T, n)$ for the Si-12 sample of Ref. [1] using the effective carrier density $n = n_s - n_c$ [5] at several values of $n_s > n_c$ and different Dingle temperatures. The impurity density, N_i, sets the overall scale of resistivity $(\rho \propto N_i)$, and does not affect the calculated T and n dependence of $\rho(T, n)$. We obtain, at low densities, both the observed nonmonotonicity and the strong drop in $\rho(T)$ in the $0.1 \sim 2$ K temperature range [1–3,9]. Our high-density results show weak monotonically increasing $\rho(T)$ with increasing T similar to experimental observations [1–3]. In the inset we show the analytic zero-temperature conductivity as a function of density n_s, following the approach of Ref. [10]. An approximately linear dependence is in good agreement with the $T \to 0$ extrapolation of the experimental [1] resistivity. Obtained results suggest that the reduced effective density and not the total value contributes to conductivity and supports our basic freeze-out or binding model [11]. These analytic results coincide with the full numerical calculation, further justifying the validity of our methods.

In conclusion, we have obtained good agreement with the experimental results. The strong temperature dependence of resistivity at low and intermediate densities $(n_s \geq n_c)$ arises from the temperature-dependent screening and a low Fermi temperature by virtue of the low effective carrier density. Thus, charged impurity scattering, carrier binding and freeze-out, temperature and density dependence of 2D screening, and classical to quantum crossover are playing significant roles in the experiments and cannot be neglected in theoretical analysis of the '2D M-I-T' phenomenon.

References

[1] S. V. Kravchenko *et al.*, Phys. Rev. **B50**, 8039 (1994); **51**, 7038 (1995); Phys. Rev. Lett **77**, 4938 (1996); D. Simonian *et al.*, *ibid* **79**, 2304 (1997); V. M. Pudalov, JETP Lett. **65**, 932 (1997); D. Popovic *et al.*, Phys. Rev. Lett. **79**, 1543 (1997).

[2] Y. Hanein *et al.*, Phys. Rev. Lett. **80**, 1288 (1998); M. Y. Simmons *et al.*, *ibid* **80**, 1292 (1998).

[3] P. T. Coleridge *et al.*, Phys. Rev. **B56**, R 12764 (1997); J. Lam *et al.*, *ibid* **56**, R 12741 (1997).

[4] V. M. Pudalov, JETP Lett. **66**, 175 (1997); V. Dobrosavljević *et al.*, Phys. Rev. Lett. **79**, 455 (1997); C. Castellani *et al.*, Phys. Rev. **B57**, 9381 (1998); S. He and X. C. Xie, Phys. Rev. Lett. **80**, 3324 (1998); D. Belitz and T. R. Kirkpatrick, Phys. Rev. **B58**, 8214 (1998); Q. Si and C. M. Varma, Phys. Rev. Lett. **81**, 4951 (1998); P. Phillips *et al.*, Nature **395**, 253 (1998); B. Altshuler and D. Maslov, Phys. Rev. Lett. **82**, 145 (1999).

[5] S. Das Sarma and E. H. Hwang, Phys. Rev. Lett. **83**, 164 (1999); T. M. Klapwijk and S. Das Sarma, Solid State Commun. **110**, 581 (1999); S. Das Sarma and E. H. Hwang, Phys. Rev. **B61**, R7842 (2000); S. Das Sarma, E. H. Hwang, and I. Žutić, in preparation.

[6] E. Abrahams *et al.*, Phys. Rev. Lett. **42**, 673 (1979).

[7] S. Das Sarma and E. H. Hwang, cond-mat/9909452, explains the parallel magnetic field effects on the quasi-2D transport behavior.

[8] T. Ando *et al.*, Rev. Mod. Phys. **54**, 437 (1982); F. Stern, Phys. Rev. Lett. **44**, 1469 (1980); F. Stern and S. Das Sarma, Solid State Electron. **28**, 158 (1985).

[9] A. P. Mills *et al.*, Phys. Rev. Lett. **83**, 2805 (1999).

[10] A. Gold and V. T. Dolgopolov, Phys. Rev. **B33**, 1076 (1986).

[11] V. T. Dolgopolov *et al.*, Phys. Rev. **B55**, R 7339 (1997); Y. Hanein *et al.*, *ibid* **58**, R 7520 (1998). See also S. Das Sarma and E. H. Hwang, cond-mat/9901117.

Energy losses of 2D electron gas due to near-surface acoustic phonon scattering

V. I. Pipa[†‡], N. Z. Vagidov[†], V. V. Mitin[†¶]

[†]*Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202, U.S.A.*
[‡]*Institute of Semiconductor Physics, Kiev, 252650, Ukraine*

M. Stroscio[§]

[§]*U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, NC 27709, U.S.A.*

We have shown that for quantum wells placed close to the stress-free surface of the semiconductor heterostructure, the energy relaxation rate of two-dimensional electrons interacting with acoustic phonons at low temperatures (Bloch–Grüneisen regime) is changed considerably in comparison with that of a two-dimensional electron gas placed in a bulk of semiconductor. The relaxation rate is enhanced in the case of a semiconductor–vacuum system and is suppressed in the case of the surface covered by a thin metal film. The enhanced energy loss is caused by additional scattering at localized and reflected acoustic waves, and the decrease appears due to suppression of piezoelectric scattering in the vicinity of the metal.

© 2000 Academic Press

Key words: Bloch–Grüneisen regime, energy relaxation rate, electron–phonon scattering, piezoelectric, deformation potentials.

1. Introduction

At low temperatures scattering with acoustic phonons is a principal process leading to energy losses of electrons in semiconductor heterostructures [1]. In recent years, the influence of modification of the acoustic phonon modes in bounded semiconductor heterostructure on the electron relaxation process has attracted substantial interest; see, e.g. [2] and references therein. It was shown that proximity of a two-dimensional electron gas (2DEG) to the surface of semi-infinite semiconductor [2–4] or slab [5] may substantially change the relaxation processes compared with that for 2DEG placed in the bulk of semiconductor. A large change in energy loss occurs at low temperatures when the phonon scattering processes are inelastic (Bloch–Grüneisen regime). This regime corresponds to temperatures T which are less than or comparable to the characteristic temperature $T_0 = 2sp_F/k_B$ where s is sound velocity, p_F the electron Fermi momentum, and k_B Boltzmann's constant. The results [2, 3, 5] were obtained for electron–acoustic phonon interaction via the deformation potential (DP). For GaAs-based heterostructures, piezoacoustic (PA) scattering dominates over that of the DP interaction [1] at low temperatures and the transition from DP-coupled to PA-coupled phonon

---

[¶]Author to whom correspondence should be addressed.

© 2000 Academic Press

scattering takes place for temperatures close to the transition temperature T_0. Thus, to study the influence of the surface on the electron kinetics in real semiconductor materials, one also has to take into account the PA scattering. To the best of our knowledge, the energy relaxation of 2DEG interacting via DP and PA potentials with the full set of phonon modes in semi-bounded systems has not been carried out to date, and that is a subject of this paper. We consider a semi-bounded heterostructure with a stress-free surface and use two types of electrical boundary conditions; one corresponds to the semiconductor-vacuum system and the other to a semiconductor covered by a thin metal film.

2. Model and basic equations

We shall consider a semiconductor heterostructure which occupies the half space $z > 0$ with a stress-free plane boundary at $z = 0$. At the distance z_0 from the surface, there is a 2D electron channel formed by the electrons in a rectangular quantum well (QW) of width d. For the sake of simplicity, we assume the same elastic and piezoelectric properties, densities, and dielectric constants for all of the layers of the heterostructure. We assume that electrons occupy the lowest subband, and that the wavefunction for transverse movement is $\psi(z) = (2/d)^{1/2}\cos(\pi(z - z_0)/d)$ for $|z - z_0| \leq d/2$ and $\psi(z) = 0$ outside the QW.

The electron system is described by the Fermi distribution function with electron temperature, T_e. The energy relaxation rate, ν_e, is introduced through the balance equation

$$\frac{2}{n_s L^2} \sum_p \varepsilon_p J_{e-ph}(\boldsymbol{p}) = -\nu_e(T_e - T). \tag{1}$$

Here n_s is the electron sheet density, L^2 is the normalization area in the xy-plane; $\boldsymbol{p} = (p_x, p_y)$, $\varepsilon_p = p^2/2m^*$ and m^* are the 2D momentum, the energy and the effective mass of electrons, $J_{e-ph}(\boldsymbol{p})$ is the integral of electron–phonon collisions. Electrons are assumed to be degenerate and the case of a small deviation from thermodynamic equilibrium is considered. The change of electron potential energy due to interaction with acoustic phonons is $V = D \operatorname{div}\boldsymbol{u} + e\phi$, where D is the deformation potential constant and ϕ is the sum of piezoelectric potential and potential induced by redistribution of the electron density. The displacement vectors, $\boldsymbol{u}$, are found from the elastic wave equation in the isotropic continuum approximation, the stress-free boundary conditions at the surface $z = 0$ are imposed. In our calculations, we exploit the full set of modes used in Refs [2,4]. In this representation, phonons are characterized by the set of quantum numbers $\omega, \boldsymbol{q}$, and j, where ω is the angular frequency, $\boldsymbol{q} = (q_x, q_y)$ is the in-plane wavevector, and the label j specifies different types of the modes. The interaction Hamiltonian is

$$\hat{H} = \sum_{j,q} \int d\omega \left(\frac{\hbar}{2\rho\omega L^2}\right)^{1/2} \left[e\phi_{j\omega q} + D\left(iqu_{j\omega q} + \frac{\partial u_{zj\omega q}}{\partial z}\right)\right] e^{i(qr-\omega t)} b_{j\omega q} + H.c., \tag{2}$$

where $b_{j\omega q}$ is a phonon annihilation operator, ρ represents the density, and $j = l, th, tv, R$. The modes l, th and tv correspond to the following choice of incident waves: longitudinal incident wave (l-mode), transverse horizontal wave polarized in xy-plane (th-mode), and transverse vertical wave polarized in the plane of incidence (tv-mode); the R-mode denotes a Rayleigh wave which has velocity s_R and obeys the dispersion law $\omega = s_R q$. For each mode, the integration is over the range where the mode exists.

The potential $\phi_{j\omega q}$ is determined from Poisson's equation. We assume that the z-axis is oriented along the (001) direction in a cubic crystal. We shall consider the two types of boundary conditions. For a semiconductor having an electrically free boundary with a vacuum, the electric potential vanishes in vacuum as the distance from the crystal increases, and the potential and normal components of dielectric displacement are continuous at the surface $z = 0$. The short-circuit condition, $\phi_{j\omega q}(z = 0) = 0$, is applied for a piezoelectric sample covered with a metal film.
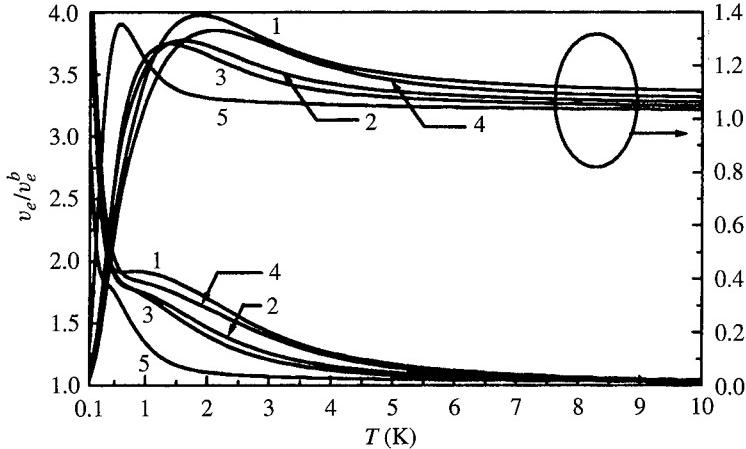
Fig. 1. Ratio of energy relaxation rates to bulk values vs. temperature for a semiconductor bounded by a vacuum (a metal film). The distance z_0 from the crystal surface to the center of a QW equals $3d/2$, where d is the width of the well. Sheet electron concentrations n_s are: 1–10^{11} cm^{-2}; 2, 4, 5–2.2 × 10^{11} cm^{-2}; 3–3 × 10^{11} cm^{-2}; well widths, d: 1, 2, 3–50 Å, 4–40 Å, 5–100 Å. The curves marked by a circle in the upper part of the figure correspond to a semiconductor in contact with a metal film.

The transition probability due to the interaction with phonons is calculated within the fermi golden rule approximation. Further derivation of the total relaxation rate may be carried out in analogy to that for DP scattering [3]; the procedure yields the following final form

$$v_e = \frac{m^{*2}}{4\pi\hbar\rho k_F^3 (k_B T)^2} \sum_j \int_0^\infty d\omega \int_0^{2k_F} dq \frac{\omega^2 |M_j(\omega, q)|^2}{|\epsilon_e|^2 \ (1 - (q/2k_F)^2)^{1/2} \sinh^2(\hbar\omega/2k_B T)}. \tag{3}$$

Hereafter, ω and q are expressed through the energy and momentum transfer of the electrons: $\hbar\omega = |\varepsilon - \varepsilon'|$, $\hbar q = |p - p'|$ in accordance with the energy and momentum conservation laws. M_j is the scattering matrix element, ϵ_e is the dielectric permittivity of 2DEG, and $k_F = p_F/\hbar$. The l, tv, and R-modes contribute to both the DP and the PA interactions, and the th-mode contributes to only the PA interaction. In the case of interest, when the surface of cubic crystal is spanned by two lattice axes, the scattering is isotropic (in xy-plane), and thus, the DP and PA mechanisms contribute to the transition rate additively: $|M_j|^2 = |M_j^{PA}|^2 + |M_j^{DP}|^2$. The dielectric permittivity ϵ_e of 2DEG depends on a QW position z_0 and is given by

$$\epsilon_e(z_0, q) = 1 + \frac{2}{a_B q} \int_0^\infty dz \psi^2(z) \int_0^\infty dz' \psi^2(z')[e^{-q|z-z'|} + g(\epsilon_0)e^{-q(z+z')}], \tag{4}$$

where $a_B = \epsilon_0 \hbar^2/m^* e^2$ is effective Bohr radius, ϵ_0 is the lattice dielectric permittivity, $g = (\epsilon_0 - 1)/(\epsilon_0 + 1)$ for contact with vacuum and $g = -1$ for contact with metal. The screening is taken into account within the Thomas–Fermi approximation which is a satisfactory one at low temperatures [2].

3. Results and discussion

Let us first discuss the temperature behaviour of the energy relaxation rates for the different types of electrical boundary conditions. For a narrow QW placed close to the surface and for the temperature T, small compared with the transition temperature T_0, the tangential and normal components of phonon wavevector, q and $q_j = (\omega^2/s_j^2 - q^2)^{1/2}$ respectively, are much smaller than the width of the well, d, and the distance z_0. The first nonzero term in the expansion of M_j^{PA} over $kz_0 \ll 1$, where $k = q_l, q_t, q$ is $(kz_0)^0$ in the case of
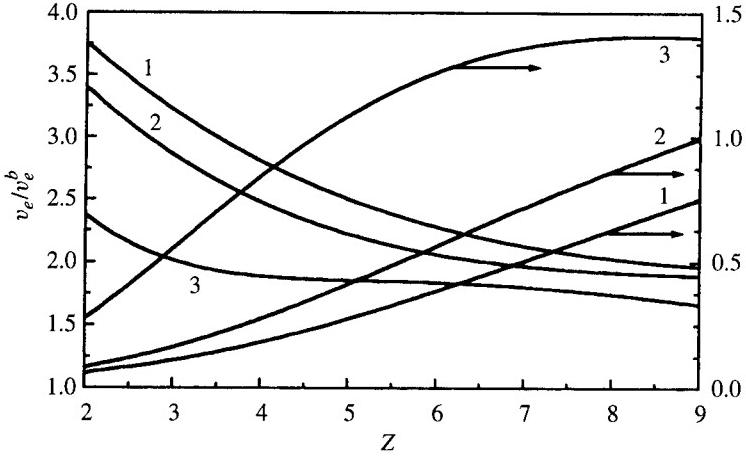
Fig. 2. Ratio of energy relaxation rates to bulk values as a function of distance from the crystal surface to the center of QW, $Z = 2z_0/d$, for a semiconductor in contact with a vacuum (a metal film) at temperature $T = 0.2$ K. The electron sheet concentrations $n_s = 10^{11}$ cm^{-2}; well widths d: 1–40 Å, 2–50 Å, 3–100 Å. The right y-axis refers to a semiconductor in contact with a metal film.
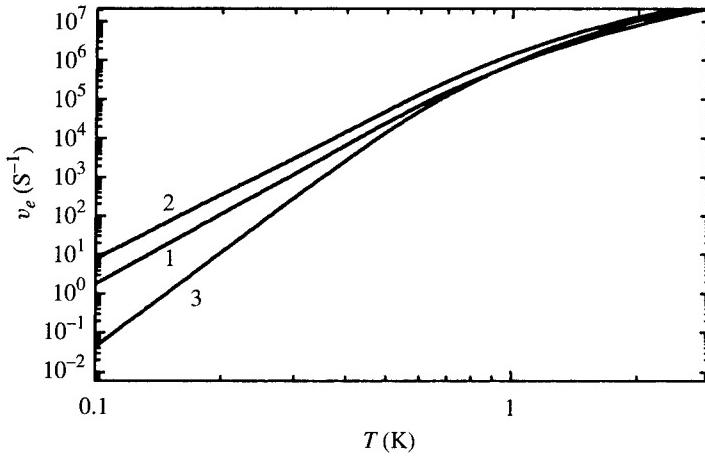
Fig. 3. Energy relaxation rates as a function of temperature for a semiconductor in contact with: 2: a vacuum, 3: metal film. Curve 1 corresponds to a bulk rate. Sheet electron concentration $n_s = 2.2 \times 10^{11}$ cm^{-2}, well width of 40 Å.

boundary with vacuum and is $(kz_0)^1$ for a surface covered by metal film. The corresponding expansion of M_j^{DP} begins with $(kz_0)^0$ regardless of a type of surface. In the limit of small q, ϵ_e increases as q^{-1} (contact with vacuum) or appears to be independent to q (metallized surface).

In the case of a semiconductor–vacuum system, from eqn (3) we get the following dependences: $v_e^{DP} \sim T^6$ for screened and $\sim T^4$ for unscreened interaction; $v_e^{PA} \sim T^4$ and $v_e^{PA} \sim T^2$, respectively. These power laws agree with known result for a 2DEG in the bulk [1]. For a semiconductor covered with a thin metal film, we find $v_e^{DP} \sim T^4$ and $v_e^{PA} \sim T^4$ for both the cases of screened and unscreened interactions. In a vicinity of metallic surface the bare PA interaction is suppressed but simultaneously the screening is decreased, so the power laws for the screened PA interaction remain the same for both types of surfaces. It should be noted that eqn (4) does not contain a small parameter which justifies the neglect the screening.

Numerical calculations were carried out for the following parameters of GaAs: $s_l = 5.2 \times 10^5$ cm s^{-1},

$s_t = 3.0 \times 10^5$ cm s^{-1}, $s_R = 2.77 \times 10^5$ cm s^{-1}, $\rho = 5.3$ g cm^{-3}, $\epsilon_0 = 12.5$, $m^* = 0.067\, m_0$, $D = 8$ eV, and the component of the piezoelectric tensor $e_{14} = 0.16$ C m^{-2}. The temperature dependences of the normalized energy relaxation rates of a 2DEG placed close to the surface of a semiconductor with vacuum and metal boundaries are shown in Fig. 1. The normalization function, ν_e^b, is the bulk value (ν_e ($z_0 \rightarrow \infty$)). Figure 1 shows that the proximity of the semiconductor–vacuum interface results in an enhancement of the electron–acoustic phonon scattering. The effect is more pronounced for thin QWs with low electron densities. As seen from Fig. 1, the temperature dependence of ν_e/ν_e^b in a semiconductor-metal system differs radically from that of a semiconductor bounding a vacuum. The most significant result illustrated in Fig. 1 is a decrease in the near-surface scattering. The rapid increase at the lowest temperatures in Fig. 1 illustrates the important role of PA scattering for a semiconductor bounded by a vacuum. At the same time, the rapid drop of the rates shown in Fig. 1 demonstrates that PA-coupled scattering near a metallized surface is suppressed. The right part of Fig. 1 shows that with the increase of temperature the relaxation rates approach to the bulk value, ν_e^b. The temperature dependences in this range are similar for the both types of boundary conditions. The enhancement of the energy loss is mainly due to additional scattering by surface-reflected and localized phonon modes. The dependences of the normalized energy relaxation rates on the QW position are shown in Fig. 2. We see that the influence of a crystal surface has a long-distance character and the corresponding scale is larger for thin electron channels. Figure 3 shows the temperature dependence of the energy relaxation rate for QW placed in the bulk of semiconductor and near the boundary with vacuum and metal film.

In conclusion, the peculiarities of the near-surface scattering originate from modification of the acoustic-phonon modes caused by the stress-free crystal surface, dependence of the phonon-induced piezoelectric potential and a dielectric permittivity of 2DEG on the dielectric properties of a medium in contact with the semiconductor.

References

[1] P. J. Price, J. Appl. Phys. **53**, 6863 (1982).
[2] Y. M. Sirenko, K. W. Kim, and M. A. Stroscio, Phys. Rev. **B56**, 15770 (1997).
[3] V. I. Pipa, F. T. Vasko, and V. V. Mitin, Phys. Stat. Solidi (b) **204**, 234 (1997); J. Appl. Phys. **85**, 2754 (1999).
[4] V. I. Pipa, N. Z. Vagidov, V. V. Mitin, and M. Stroscio, Physica **B270**, 280 (1999).
[5] V. I. Pipa, B. A. Glavin, V. V. Mitin, and M. Stroscio, Semicond. Sci. Technol. **13**, A97 (1998).

Electron-induced dissociation of SiH complexes in hydrogenated Si-doped GaAs. Application to the fabrication of microstructures

S. SILVESTRE, E. CONSTANT, D. BERNARD-LORIDANT

I.E.M.N., UMR CNRS 9929, BP 69, Avenue Poincaré, 59652 Villeneuve d'Ascq Cedex, France

M. CONSTANT

L.A.S.I.R., UMR CNRS 8516, USTL, 59655 Villeneuve d'Ascq Cedex, France

J. CHEVALLIER

L.P.S.C., UMR CNRS 8635, 1, Place Aristide Briand, 92195 Meudon Cedex, France

We study the role of hot electron injection in the dissociation of the SiH complexes which appear in n-type Si-doped GaAs epilayer exposed to a hydrogen or deuterium plasma. Firstly, the results recently obtained in room-temperature aging experiments on hydrogenated or deuterated Schottky diodes submitted to high bias voltage are summarized and the role of hot carriers in the dissociation of donors and the observed isotope effect are described. Then, it is shown that SiH dissociation can also be achieved using hot carriers injected into the semiconductor by electron beam. Such electron-beam effects are finally used for the fabrication and characterization by cathodoluminescence of micronic conductive GaAs structures.

© 2000 Academic Press

Key words: hydrogen, III–V semiconductors, isotope effect, energetic electrons, nano-structures.

1. Introduction

In the last few years, considerable effort has been focused on the role of hydrogen in crystalline semiconductors and on its influence on the reliability of Si or III–V components. For Si-based microelectronic devices, studies have been devoted to metal-oxide-semiconductor (MOS) Si transistors. In such devices, hot electron degradation due to the breaking of SiH bonds and to the formation of dangling bonds is observed and recently, it has been shown [1] that the reduction of this degradation and the improvement of the device lifetime could be achieved by incorporating deuterium rather than hydrogen. For Si-doped GaAs materials and devices, the ability of hydrogen to passivate the electrically active shallow impurities and as a consequence to form an electrically neutral SiH complex is well known [2]. So, it is of interest to study the influence of hot electrons on the stability of complexes and the related isotope effects. Here, we will present in the first part of this paper the main results of an experimental study [3] regarding the complex dissociation occurring in hydrogenated or deuterated reverse bias GaAs Schottky diodes where hot electrons are present. It has been

demonstrated, on a hydrogenated Si surface, that external energetic electrons injected from a scanning tunneling microscope (STM) tip induce a breaking of SiH bonds [4]. This phenomenon has already successfully been applied for the realization of mesoscopic structures in silicon [5]. Here, we present results on related effects in hydrogenated Si-doped GaAs. The reactivation of neutralized dopants using external electrons injected by an electron-beam lithography system is shown and the application of this process to the fabrication of conductive microstructures is demonstrated.

2. Room-temperature aging experiments: a strong isotope effect

These experiments have been carried out on Schottky diodes fabricated from 1 μm of Si-doped GaAs ($n = 2.2 \times 10^{17}$ cm^{-3}) epilayers grown on a n$^+$-GaAs substrate by molecular-beam epitaxy. The wafers were first hydrogenated or deuterated in a 13.56 MHz hydrogen (H) or deuterium (D) plasma under the following conditions: temperature of 190 °C, gas pressure of 1 mbar, RF power of 0.16 W cm^{-2}, for 15 min for hydrogen and 25 min for deuterium. These plasma exposure times have been selected in order to give the same dopant passivation depths for H and D atoms. Schottky contacts were then placed on the hydrogenated surface. With such a structure, the depth distribution of the active (i.e. ionized) donor concentration can be easily determined from capacitance–voltage measurements with a HP4194 RF impedance analyser.

Figure 1 shows the results obtained on representative samples after plasma exposure and after temperature annealings or bias stress aging experiments. Due to the selected exposure time, H and D plasma have passivated the donors in a very similar way but for the legibility of the figure, only the result of the hydrogenated diode before annealing and bias stress experiments is represented. Due to the SiH complex formation, the active doping concentration has been reduced from 2.2×10^{17} cm^{-3} to a few 10^{16} cm^{-3} over the first 0.22 μm. In order to study the stability of complexes, various aging or annealing experiments have then been carried out for temperatures above 170 °C or bias greater than 4 V, which are critical values to observe significant changes. The first example is an annealing experiment performed on hydrogenated or deuterated samples without bias voltage at 250 °C for 1 h. It should be noted that the complex dissociation occurs quite uniformly in all the passivated region and that the thermal dissociation rates for hydrogenated and deuterated samples are close to each other. Concurrently, aging experiments have been performed at 20 °C for a constant bias voltage of 5.5 V and two examples of the results obtained on hydrogenated and deuterated diodes are represented in Fig. 1. In the case of the deuterated sample, it can be noted that the complex dissociation only occurs at the extremity of the deuterated region. In this region, the 'lucky electron' which has been injected through the metal–semiconductor interface and which has not suffer of too many inelastic collisions might acquire the highest energy value and this consideration suggests that the SiH complex dissociation could be due to these hot electrons. Consequently, in order to compare the results obtained on hydrogenated and deuterated diodes, the stress time for each experiment has been selected monitoring the Schottky diode reverse current in order to have the same number of electrons injected in the passivated region. Then, it clearly appears that dissociation effects are much less important in deuterated diodes than in hydrogenated samples where the dissociation effects have occurred in a much wider part of the passivated region. Taking into account all these results, it appears that hot electron excitations seem to play an important role in the dissociation of SiH complexes and that the involved physical phenomena are associated with a strong isotope effect.

3. Energetic electron-induced reactivation

In order to obtain more insight into the role of electrons in the dissociation of SiH complexes, we can now study the case where external energetic electrons are injected into the material. This can be achieved by using scanning tunneling microscopy (STM) tips but the simplest and easiest method for possible future applications in the fabrication of nanodevices seems to be the use of an electron beam produced by a
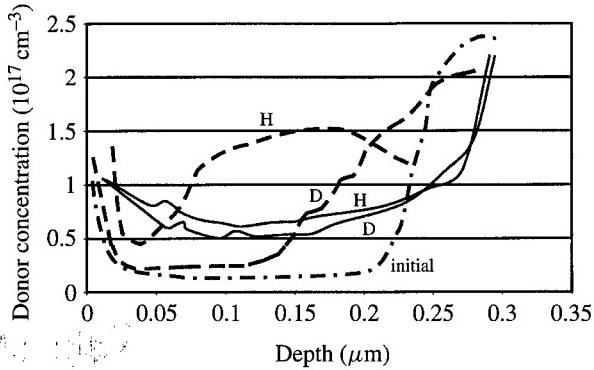
Fig. 1. $C-V$ depth profiles of the active Si donor concentration for hydrogenated (H) or deuterated (D) Schottky diodes. –•–•, Hydrogenated diode after plasma exposure and before annealing or bias stress (the results obtained for deuterated samples are very close to those obtained for hydrogenated samples). —, Hydrogenated or deuterated diodes simultaneously annealed without bias at 250 °C for 1 h. – –, Hydrogenated or deuterated diodes submitted at room temperature to bias stress at a reverse voltage of 5.5 V. The bias stress time (about 5 h) have been selected for H and D samples to have the same given electron number $N_e = 16.7 \times 10^{18}$ injected in the diodes.
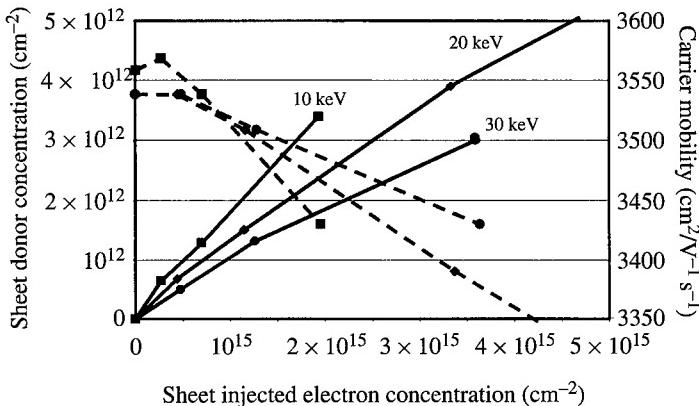
Fig. 2. Evolution of the sheet carrier concentration (full line) and of the electron mobility (dashed line) measured by the Hall effect on hydrogenated n-type Si-doped GaAs epilayer versus the electron number dose for various electron injection energies : 10 (■), 20 (♦) and 30 keV (●).

scanning electronic microscope (SEM) or an electron lithography system. Such a method has already been used for acceptors reactivation in Mg-doped GaN [6]. To test this possibility in the case of Si- doped GaAs, experiments have been carried out on 0.35 μm thick Si-doped epilayers grown by molecular-beam epitaxy. From the Hall measurement, sheet electron concentration $N_s = 1.4 \times 10^{14}$ cm^{-2} and free carrier mobility $\mu n = 1690$ cm^2 V s^{-1} have been obtained. The layers are then exposed to a RF plasma under the following conditions: temperature of 190 °C, hydrogen pressure of 1 mbar, RF power of 0.16 W cm^{-2} for 8 h. After plasma exposure, the samples are characterized for $N_s = 4.6 \times 10^{12}$ cm^{-2} and $\mu n = 3550$ cm^2 V s^{-1}. These results can be explained by the neutralization of ionized donors, and the increase of mobility is due to the decrease of ionized impurity scatterings. Then, we have exposed the whole clover surface to the focused electron beam of a SEM with different electron injection energies: 10, 20 and 30 keV. The obtained results for an increasing incident electron number are presented in Fig. 2.

20 μm
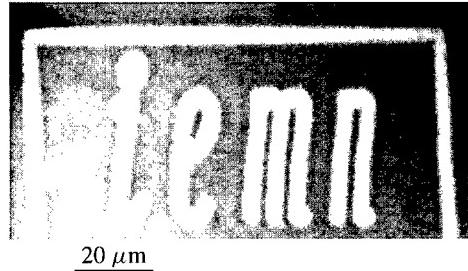
Fig. 3. CL image of the laboratory logo written with a 20 keV electron beam on hydrogenated Si-doped GaAs sample. The electron dose is 0.5 C cm^{-2}. The CL reading of the patterns has been made on a sample cooled down to liquid nitrogen and the CL signal has been collected by an ellipsoidal mirror of an Oxford system mounted on the scanning microscope.

The following can be noted: (i) a strong increase of the free carrier concentration and a decrease of the electron mobility when increasing the number of electrons injected in the material; and (ii) an higher efficiency of electron-induced dopant reactivation as the electron injection energy decreases. The first phenomenon can be explained by the dissociation of SiH complexes due to highly energetic electrons and an increase of the ionized impurity scattering, respectively. The second observation can be explained by the well-known results of electron trajectories Monte Carlo simulations which clearly show that the maximum energy loss of electrons occurs for larger penetration depths when the electron energy increases. Actually, if we consider 10 keV acceleration voltages, the penetration depth of electrons is 0.5 μm. Most of its interactions take place within the thickness of the first 0.35 μm Si-doped layer, and as a result, numerous interactions between electron and SiH complexes might occur. However, if we now study the interaction of 30 keV accelerated electrons, the penetration depth is 3.3 μm, and thus most of its interactions occur in the semiconductor substrate below the active layer where SiH complexes are not present. These considerations explain our observations for different electron injection energies. It is pointed out that exposing hydrogenated n-type GaAs:Si epilayers could allow the reactivation of free carriers very locally. So, it should be possible, starting from a quasi-insulating hydrogenated n$^+$ epilayer, to realize variable shape patterns of micronic or nanometric regions characterized by a high electron concentration. In order to exploit this possibility, we used cathodoluminescence (CL) imaging based on the increase of the luminescence intensity area according with the active doping level. In a first experiment we draw our laboratory's logo on a hydrogenated epilayer with a 20 keV energy electron beam stemmed from an electron-beam lithography system. The size of the letters have been chosen according to spatial resolution imaging of the CL, which is about 1 μm. However, electron beam permits the realization of more nanometric-scaled structures. The results obtained are presented in Fig. 3.

The dimensions of the letter which are observed in the CL images are very slightly larger than the electron-beam lithography ones but the difference observed are below the CL spatial resolution. Consequently, other type of reading method will have to be used to characterize, at a more nanometer scale, the mesoscopic structures that could be fabricated by electron-beam writing in hydrogenated GaAs.

4. Conclusion

It has been shown that the reactivation of neutralized dopants in hydrogenated n-type GaAs doped with silicon can be induced either by hot electrons produced into material in Schottky diodes (with, in this case, the observation of a strong isotope effect), or by energetic electrons issued from an electron-beam system. Such effects could be due to an electronic excitation of the SiH bonds lowering the dissociation barrier height [7]. The excitation energy is then close to 4 eV and such an energy can be provided by the hot electrons produced or injected in the GaAs : Si, H. In the case of electrons issued from an electron-beam system, our results show

a novel method of fabrication of conductive GaAs microstructures. However, to fully take advantage of this technique to fabricate nanostructures, a suitable way to read what has been fabricated has to be found (for example by using STM). In addition, it should be noticed that electron-beam-induced reactivations should also occur in hydrogenated heteroepitaxy with a 2D electron gas and such an effect could open the fabrication of high mobility 1D or 2D mesoscopic structures.

References

[1] J. W. Lyding, K. Hess, and I. C. Kisilyalli, Appl. Phys. Lett. **68**, 2526 (1996).

[2] J. Chevallier, W. C. Dautremont-Smith, C. W. Tu, and S. J. Pearton, Appl. Phys. Lett. **47**, 108 (1985).

[3] E. Constant, D. Bernard-Loridant, S. Mezière, M. Constant, and J. Chevallier, J. Appl. Phys. **85**, 6526 (1999).

[4] Ph. Avouris, R. E. Walkup, A. R. Rossi, T.-C. Shen, G. C. Abeln, J. R. Tucker, and J. W. Lyding, Chem. Phys. Lett. **257**, 148 (1996).

[5] C. Syrykh, J.-P. Nys, B. Legrand, and D. Stiévenard, J. Appl. Phys. **85**, 3887 (1999).

[6] H. Amano, M. Kito, K. Hiramatsu, and I. Akasaki, Jpn. J. Appl. Phys. **28**, L2112 (1989).

[7] Y. Miyamoto, O. Sugino, and Y. Mochizuki, Appl. Phys. Lett. **75**, 2915 (1999).

Anharmonic linewidth of absorption by localized vibrations of H and D adatoms on the surface of Si

I. P. IPATOVA, O. P. CHIKALOVA-LUZINA

A.F.Ioffe Physical Technical Institute, Russian Academy of Sciences, 194021, St.-Petersburg, Russia

K. HESS

Beckman Institute, The University of Illinois, Urbana, U.S.A.

Measurements of the integrated IR absorption by stretching localized vibrations of H adatoms on the Si (111) surface [1] have shown an anomalous temperature dependence. The line intensity (integrated area) decreases by 20% as the temperature is raised from 130 to 560 K. The authors of [1] believe that the anomalous temperature dependence follows from a strong anharmonic interaction of the H–Si stretching modes with H–Si bending modes.

All the facts found for this case are closely related to the widely discussed optical properties of U-center localized vibrations in alkali halides. It has been shown by Hughes [2] and Ipatova *et al.* [3] that the anomalous temperature dependence the integrated intensity results from the large difference of the localized frequency and the frequencies of the bulk modes of alkali halides (adiabatic approximation). It has been shown that the strength of the absorption by the localized mode has an exponential temperature dependence of the Debye–Waller type.

This paper deals with the theory of the anharmonic linewidth and integrated absorption by localized vibrations of H- or D-adatoms on the Si (111) surface. The anharmonic interaction of the Si–H stretching localized mode with the substrate vibrational Si modes is taken into account. Since the Si–H stretching mode frequency is $\omega_{loc}(H) = 2085 \, \text{cm}^{-1}$, the stretching mode frequency of Si–D is $1516 \, \text{cm}^{-1}$ [4], and the maximum phonon frequency of Si is $\omega_L(\text{Si}) = 514 \, \text{cm}^{-1}$, the adiabatic parameter,

$$\lambda = \frac{\omega_L}{\omega_{loc}} < 1, \tag{1}$$

holds for both H–Si and D–Si adatoms. This means that due to the anharmonic interaction, the slow motion of Si-atoms occurs in an averaged field of the light adatoms fast vibrations. It will be shown that there appear static displacements of the Si atom equilibrium sites and vibrations of the Si atoms occur about new equilibrium positions.

The shift of equilibrium positions results in an exponential temperature-dependent factor (Debye–Waller factor) for both the integrated absorption and the spectral linewidth of the light absorption by localized vibrations of the H or D adatoms. This exponential factor depends on the temperature and on the mass of the adatom. It is different for the localized vibrations of the isotope complexes Si–H, Si–D.

We assume that the adatom (H or D) of mass M' is located on the (111) surface of Si. The adatom is supposed to be bound by a force characterized by the 'on-site' configuration to the surface atom of Si at the lattice site $l = (l_x \, l_y \, l_z) = (0\,0\,0) \equiv 0$. The position of adatom is $l = (0\,0\,1) \equiv 1$. The mass of the Si atom is M.

The coefficient of the absorption of IR light at frequencies $\omega \approx \omega_{\text{loc}}$ has the form [5]

$$K(\omega) = -c_d \frac{e^2}{\hbar v_0} \operatorname{Im} D_{\alpha\beta}^R(1, 1; \omega) \frac{E_\alpha E_\beta}{E^2}, \tag{2}$$

c_d being the surface concentration of adatoms, v_0 is the volume of the crystal primitive cell and $D_{\alpha\beta}^R(1, 1; \omega)$ is the one-particle retarded Green's function of the adatom localized vibrations.

The retarded Green's function $D_{\alpha\beta}^R(1, 1; \omega)$ can be calculated by the temperature Green's function technique. The temperature Green's function $D_{\alpha\beta}(1, 1; i\omega_n)$ with $\omega_n = 4\pi nT/\hbar$ satisfies the Dyson equation

$$D_{\alpha\beta}(1, 1; i\omega_n) = D_{\alpha\beta}^{(0)}(1, 1; i\omega_n) + \sum_{ls, l's'} D_{\alpha\gamma}^{(0)}(1, ls; i\omega_n) \Pi_{\gamma\delta}(ls, l's'; i\omega_n) D_{\delta\beta}(l's', 1; i\omega_n), \tag{3}$$

where $ls, l's'$ are positions of the crystal lattice sites; $\Pi_{\gamma\delta}(ls, l's'; i\omega_n)$ is the polarization operator produced by the anharmonic terms in the potential energy of the system. Taking into account the symmetry of the adatom position on the surface, one can find $D_{\alpha\beta}(1, 1; i\omega_n) = D(1, 1; i\omega_n)\delta_{\alpha\beta}$ and $\Pi_{\alpha\beta}(1, 1; i\omega_n) = \Pi(1, 1; i\omega_n)\delta_{\alpha\beta}$.

The adiabatic parameter, eqn (1), allows us to neglect the nondiagonal terms $D(1, ls; i\omega_n)$ in the sum over $(ls, l's')$ in eqn (3) and to find $D(1, 1; i\omega_n)$ algebraically. The retarded Green's function $D^R(1, 1; \omega)$ is the analytic continuation in ω of the temperature Green's function $D(1, 1; i\omega_n)$. One can separate the real and the imaginary parts of $\Pi(1, 1; \omega)$ to obtain $\Pi(1, 1; \omega) = \Delta(1, 1; \omega) - i\Gamma(1, 1; \omega)$, which determine the anharmonic shift $\operatorname{Re} \Pi \equiv \Delta(1, 1; \omega)$ of the localized frequency and the decay of the localized vibration $\operatorname{Im} \Pi \equiv \Gamma(1, 1; \omega)$.

The quantity $\Gamma(1, 1; \omega)$ is defined by the processes which satisfy the conservation laws of energy and momentum. When the third-order anharmonic term is taken into account, in the first approximation of the perturbation expansion, the term $\Gamma(1, 1; \omega_{\text{loc}})$ does not vanish when $\omega_L < \omega_{\text{loc}} < 2\omega_L$, a condition which holds neither for H- nor for D-adatoms. In order to find $\Gamma(1, 1; \omega)$ the anharmonic terms of the fourth order should also be taken into account. There are no limitations of this type for the shift $\Delta(1, 1; \omega)$: all possible virtual anharmonic phonon processes contribute to Δ. As a result, $\Delta(\omega) > \Gamma(\omega)$ and the shift Δ are defined by the third order anharmonisity. The expansion of $\Pi(1, 1; \omega)$ near $\omega \approx \omega_{\text{loc}}$ gives

$$D^R(1, 1; \omega) = \frac{1}{1 - \Delta'(\omega_{\text{loc}})} \left(\frac{\hbar}{M'} \right) \frac{1}{\omega - \omega_c(T) + i\gamma(\omega)}, \tag{4}$$

where

$$\omega_c^2(T) = \omega_{\text{loc}}^2 + \frac{\Delta(\omega_{\text{loc}})}{1 - \Delta'(\omega_{\text{loc}})} \qquad \text{and} \qquad \gamma(T) = \frac{\Gamma(\omega_{\text{loc}})}{1 - \Delta'(\omega_{\text{loc}})}. \tag{5}$$

Here $\Delta'(\omega_{\text{loc}}) = \left(\frac{d\Delta}{d\omega^2} \right)_{\omega^2 = \omega_{\text{loc}}^2}$.

Calculations of the polarization operator $\Pi(1, 1; \omega)$ and the corresponding ω_c and γ using the temperature Green's function technique allow us to find the main contribution to $K(\omega)$ using eqn (1).

$$K(\omega) = c_d \frac{e^2}{\hbar v_0} \left(\frac{\hbar}{2M'\omega_{\text{loc}}} \right) e^{-2W(T)} \frac{e^{-2W(T)}\Gamma}{(\omega - \omega_c)^2 + (e^{-2W(T)}\Gamma)^2}. \tag{6}$$

Here the quantity W(T) is defined by the third order anharmonic terms $\Phi^{(3)}$

$$2W(T) = \sum_\lambda \frac{\hbar}{2M\omega_\lambda} \left(\frac{\hbar}{2M'\omega_{\text{loc}}} \right)^2 |\Phi^{(3)}|^2 \frac{(2N_\lambda + 1)}{(\hbar\omega_\lambda)^2}, \tag{7}$$

where ω_λ, N_λ are the normal mode frequencies and the occupation phonon numbers of vibrations in the semi-infinite Si and $\Phi^{(3)}$ is the closure of the anharmonic force constants with the phonon polarization vectors.

It follows from eqn (6) that the quantity $e^{-2W(T)} \Gamma(\omega_{\text{loc}})$ represents the linewidth of the optical absorption by the high-frequency localized mode. At low temperatures $T < \Theta_D$, Θ_D being Debye temperature, the factor e^{-2W} does not depend on temperature: $e^{-2W(0)} = \text{const}$. At high temperatures $T > \Theta_D$ the Debye–Waller factor is a linear function of temperature.

Two contributions to the quantity Γ are known: the relaxation, Γ_1, and the dephasing, Γ_2. At low temperatures $T < \Theta_D$, the main contribution to the line width for an H adatom $\Gamma(H)$ arises from the dephasing mechanism of the broadening. $\Gamma_2(H)$ vanishes because $e^{-\frac{\hbar\omega_\lambda}{T}}$ at $T \to 0$. In case of D-adatoms, the dephasing Γ_2 defines the linewidth in the temperature interval $T > T_1$ only. At $T = T_1 \approx 60$ K. the relaxation broadening $\Gamma_1(D)$ becomes more important.

At low temperatures $T < \Theta_D$, the comparison of the line width for H- and D-adatoms shows that at $T > T_2 \approx 30$ K the main mechanism of broadening is dephasing Γ_2 for both H- and D-adatoms. Order of magnitude estimates show that $\Gamma_2(H) > \Gamma_2(D)$. At $T \leq T_2$ the relaxation broadening $\Gamma_1(D)$ becomes more important for the D-line width and the H-linewidth becomes smaller than the D-line width: $\Gamma_2(H) < \Gamma_1(D)$. At high temperatures $T > \Theta_D$, the temperature dependence of the broadening is defined by $\Gamma(T) e^{-2W(T)}$ where $\Gamma(T)$ is the dephasing process for both H- and D-adatoms. Numerical evaluations show that at $T \geq 250$ K, the relaxation between the linewidths, $\Gamma(D) e^{-2W(D)} > \Gamma(H) e^{-2W(H)}$, holds due to the exponential factor.

The integrated absorption is equal to

$$J = \int_0^\infty d\omega K(\omega) = c_d \frac{e^2}{h v_0} \left(\frac{\hbar}{2M' \omega_{\text{loc}}} \right) e^{-2W(T)}. \tag{8}$$

It follows from eqn (8) that the integrated intensity J has an exponential temperature dependence which is stronger in case of H-adatoms than for D-adatoms. An order of magnitude evaluation has shown that $J(H) \approx 2 J(D)$ in good agreement with experimental data from [6] where $J(H) = 1, 9 J(D)$ is obtained.

Note that the linewidth of the IR integrated intensity is closely related to the probability of hydrogen and deuterium desorption from silicon surfaces which has recently received much attention because it determines, to a large extent, transistor lifetimes for the metal-oxide-semiconductor system. The above results hint towards a faster desorption rate for hydrogen and therefore to a reduced transistor lifetime as compared with deuterium-treated transistors as observed experimentally. The theory presented in this paper therefore represents a first step towards an analytical understanding of transistor reliability as determined by H/D and desorption processes.

References

[1] P. Dumas, Y. J. Chabal, and G. S. Higashi, Phys. Rev. Lett. **65**, 1124 (1990).

[2] A. E. Hughes, Phys. Rev. **173**, 860 (1968).

[3] I. P. Ipatova, A. V. Subashiev, and A. A. Maradudin, Ann. Phys. **53**, 376 (1969).

[4] V. A. Burrows, Y. J. Chabal, G. S. Higashi, K. Raghavachari, and S. B. Christman, Appl. Phys. Lett. **53**, 998 (1988).

[5] I. P. Ipatova and A. A. Klotchichin, Sov. Phys.-JETP **23**, 1068 (1966).

[6] H. Luo and C. E. D. Chidsey, Appl. Phys. Lett. **72**, 477 (1998).

Hydrogen-related defect creation at the Si(100)–SiO$_2$ interface of metal-oxide-semiconductor field effect transistors during hot electron stress

B. R. TUTTLE, K. HESS, L. F. REGISTER

Beckman Institute, University of Illinois, Urbana, IL 61801, U.S.A.

We explore the hydrogen-related microstructures involved in hot electron defect creation at the Si(100)–SiO$_2$ interface of metal-oxide-semiconductor field effect transistors. Based on the energetics of hydrogen desorption from the interface between silicon and silicon-dioxide, we argue that the hard threshold for hydrogen-related degradation may be considerably lower than the previously assumed value of 3.6 eV. Also, hydrogen atoms released from Si–H bonds at the interface by hot electron stress are trapped in bulk silicon near the interface.

© 2000 Academic Press

Key words: DFT, theory of MOSFETs, point defects in semiconductors.

1. Introduction

In ubiquitous metal-oxide-semiconductor field effect transistors (MOSFETs), hydrogen is used primarily to passivate interfacial defects that occur because of the lattice mismatch between the oxide (SiO$_2$) and the semiconductor (Si). From hydrogen–deuterium isotope experiments [1], it has recently been shown that hot electron degradation in MOSFETs is due in large part to the de-passivation of Si–H(D) bonds at the Si–SiO$_2$ interface. These results have been verified by both academic [2] and corporate research [3] teams. The degradation in MOSFETs has been argued to be limited by hydrogen diffusion away from the interface into bulk SiO$_2$. [4, 5] Using charge pumping and capacitence-voltage experiments, Devine *et al.* [2] and Cartier *et al.* [6] determined the charge transition energies for defects at the Si–SiO$_2$ interface of MOSFETs. These experiments place stringent limits on the possible atomic models for hot electron degradation. Theoretically, several groups [7, 8, 10, 11], including our own [9, 12–14], have investigated with *ab initio* density functional calculations the energetics and defect levels of hydrogen in Si, SiO$_2$ and at their interfaces. In the present paper, we analyse the results of theory and experiment to develop an understanding of the microscopic behavior of hydrogen during hot electron stress in MOSFETs.

The conclusions of our analysis are as follows. Based on the energetics of hydrogen desorption from the interface between silicon and silicon-dioxide, we argue that (i) the hard threshold for hydrogen-related degradation may be considerably lower than the previously assumed value of 3.6 eV, and (ii) hydrogen atoms released from interfacial Si–H bonds by hot electron stress are trapped in bulk silicon near the interface.
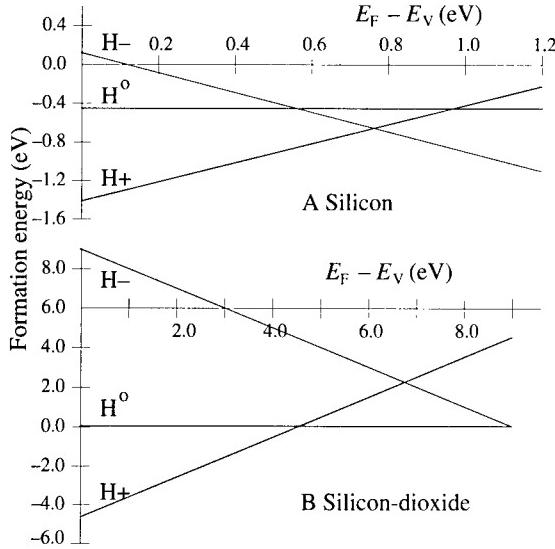
Fig. 1. Hydrogen formation energies as a function of the electronic Fermi energy of A, silicon and B, silicon-dioxide. The valence band maximum determines the zero of energy for both plots. Comparing A, with B, we note that neutral hydrogen will favor the silicon over the silicon-dioxide interface. In A, notice that hydrogen can lower its formation energy by becoming charged and the transition from positively to negatively charged is just below 0.8 eV. This transition for bulk silicon and is also observed at the Si–SiO$_2$ interface of MOSFETs during hot electron stress.

2. Theory of hydrogen in silicon and silicon-dioxide

Previously, calculations based on density functional theory within the local density approximation (LDA) [15] have been used to examine hydrogen and silicon dangling bonds in bulk Si [7–9], in bulk SiO$_2$ [10–12] and at their interface [13, 14]. In Fig. 1 we report our current estimates for the formation energy [14] of atomic hydrogen in (A) silicon and (B) silicon-dioxide. The results in Fig. 1 involve six calculations in total. For both Si and SiO$_2$, we have calculated the formation energy of hydrogen for each charge state in its minimum energy configuration. We have applied an *a posteriori* shift of the defect levels to correct for known deficiencies in the LDA, i.e. the LDA tends to under-estimate semiconductor bandgaps [16]. Therefore the total energies have been shifted up by a value equal to the defect's charge times the LDA bandgap error. We have used a 0.5 and 3.0 eV correction for silicon and silicon-dioxide, respectively [14]. Using this correction for hydrogen in bulk silicon results in theoretical donor and acceptor levels which agree nicely with experiment [17]. For H in SiO$_2$ there are no experiments to compare with. For full details of the calculations see Ref. [14]. Although calculations including the strain at the interface has not been performed, the bulk calculations reported in Fig. 1 provide important insight.

Figure 1 reports the formation energy as a function of Fermi energy within the respective bandgap regions for H^q where q is the charge state which can be $+1$, 0 and -1. Comparing Fig. 1A with 1B we notice that neutral hydrogen in silicon has a formation energy about 0.4 eV lower than in silicon-dioxide. Therefore, neutral hydrogen will favor bulk silicon over bulk silicon-dioxide. From Fig. 1A, one can see that the neutral charge state is not favored for any value of the Fermi energy in silicon. Therefore, if hydrogen has time to exchange charge with the Fermi energy in silicon, then the hydrogen will either be positive or negative. To compare charge states of hydrogen in silicon and silicon-dioxide, one needs to know how to align the bands. For our present purposes, we use the experimental valence band offset of 4.1 eV between SiO$_2$ and Si [18]. Given the 4.1 eV band offset, the formation energy of H$^+$ is over 1 eV lower in silicon than in silicon-

dioxide. These results indicate that hydrogen favors the bulk silicon environment which can be qualitatively understood because SiO_2 is a wide bandgap insulator with a strong Si–O bond strength.

As pointed out above, the donor and acceptor levels derived from Fig. 1A for hydrogen in silicon agree with experiment nicely (to within ± 0.1 eV). In silicon the hydrogen will change from positively to negatively charged at a Fermi energy of just less than 0.8 eV or about 0.2 eV above mid-gap. This is an important signature of atomic hydrogen in a bulk silicon environment.

Hydrogen is introduced into the interface in order to passivate silicon dangling bonds which occur on the silicon side of the interface. The dangling bond changes from positive to neutral and neutral to negative at 0.3 and 0.8 eV, respectively, above the valence band edge. The bond strength of the Si–H bond formed by H passivation is about 3.6 eV (the bond strength for isolated Si–H bonds in c–Si). To compare with Fig. 1A the Si–H formation energy is -3.6 eV.

3. Magnitude of the threshold energy of hydrogen-related damage

The Si–H binding energy (3.6 eV) is commonly assumed to be the threshold energy for H-related degradation in MOSFETs [4]. However, such a threshold appears inconsistent with experimental findings. If we were to assume for the moment that a hard threshold of 3.6 eV for hydrogen desorption, it would be difficult to explain degradation for bias voltages below this level. The electron energy distribution is understood to decrease *exponentially* above the energy corresponding to the operating voltage according to Boltzmann's law with T taken as the lattice temperature. Therefore, one would predict large differences in degradation as the operating voltage is scaled down. Lowering the operating voltages from 3.3 to 2.5 V should lead to a factor of 10^{20} reduction in degradation. This is not observed in practice which has led some to suggest that an enhancement of the tail of the carrier distribution via carrier–carrier interactions could explain the continuing degradation of low-bias devices [19, 20].

Previously, we have pointed to alternative explanations for the continuing degradation at low bias voltages that do not rely on overly excited carrier populations [22]. It may be that the barrier to desorption is lower than previously thought, and/or that all the energy required to overcome the desorption barrier need not be supplied by any single carrier. For example, recent investigations of the hydrogen–deuterium isotope effect in MOSFETs confirm that an important mechanism of hot-carrier degradation is indeed the breaking of Si–H(D) bonds at the interface [1]. However, studies of Si–H(D) desorption with scanning tunneling microscopy (STM) have shown that carriers with energies less than 3.0 eV can result in bond breaking.

We now turn to the energetics of silicon–hydrogen (Si–H) bonding and dissociation at the Si–SiO_2 interface. For a Si–H bond at the Si–SiO_2 interface, if the dissociated hydrogen atom enters bulk SiO_2, then the dissociation or desorption energy is 3.6 eV because atomic hydrogen interacts only weakly with the rather open, insulating oxide. However, the Si–H desorption energy can be significantly reduced for Si–H bonds at the Si–SiO_2 because hydrogen can desorb by first entering bulk silicon. The energy needed to place a neutral hydrogen atom, arising from the silicon dangling bond site, into bulk silicon far from any defects is 3.0 eV. This energy may actually be lower by several tenths of an electron volt because both the hydrogen atom and the silicon dangling bond can lower their energy by becoming charged species. The barrier for atomic hydrogen to move along bulk interstitial sites is about 0.5 eV [8]. As hydrogen diffuses to a surface or interface, it can passivate other defects or combine with another hydrogen atom to form H_2. At a surface or an open interface such as the Si–SiO_2 interface, H_2 molecules can easily diffuse away leaving behind the silicon dangling bonds. Brower found thermally activated desorption of hydrogen from the (111)Si–SiO_2 interface is measured at 2.56 eV [21] consistent with the calculated mechanism with H entering bulk silicon before leaving the system as H_2. Also, in the presence of free carriers, e.g. during MOSFET operation, the desorption barrier can be lowered to 1.5–2.0 eV [22].

In addition to the above considerations, the threshold energy can be greatly reduced if desorption occurs by multiple vibrational excitations. For low voltages, Si–H dissociation involving multiple vibrational excitation

by the transport electrons becomes relatively more likely. Because hydrogen is very light, the hydrogen in a Si–H bond is a quantum oscillator. Hot electrons can excite the hydrogen quantum oscillator from the ground state into an excited state. Since the Si–H bending and stretching vibrational modes are well above the silicon phonon modes, the excited state will be long-lived which allows for multiple vibrational excitation[7]. In this case, each channel electron needs the vibrational excitation energy of the Si–H bond which is less than 0.3 eV [8]. In general, we expect the threshold energy for hydrogen-related damage to be considerably less than 3.6 eV.

4. Hydrogen defects created during hot electron stress

Now consider the electronic defects created by Si–H dissociation during hot electron stress. The isotope effect [1–3] indicates that hydrogen plays a significant role in hot electron degradation of MOSFETs. With charge pumping measurements in 0.25 micron MOSFETs [2], the density of interface traps has been determined as a function of silicon Fermi energy position and hot electron stress. After hot electron stress, a distributions emerged with two peaks at 0.25 eV *below* and 0.2 eV *above* silicon mid-gap. The lower peak was at 1×10^{11} eV^{-1} cm^{-2} and can be associated with the $(+/0)$ transition of the silicon dangling bond defects. The total upper distribution has a peak maximum of six times the lower peak maximum whereas experiments indicate that a peak maximum of, at most, only twice as high would be expected for the $(0/-)$ transition of the silicon dangling bonds alone [23]. The data can be explained if for every dangling bond created, a hydrogen atom is trapped near the interface. For atomic hydrogen trapped on the silicon side of the interface there is a $(+/-)$ transition at 0.2 eV above mid-gap. The addition of the hydrogen defect state in silicon to the interfacial dangling bond state qualitatively explains the charge pumping experiments of Ref. [2]. Our LDA calculations suggest that, near the Fermi energies of interest, only the $(+/0)$ transition is possible for atomic hydrogen in SiO$_2$, see Fig. 1B. This would lead to a peak lower than observed. Also, if hydrogen were to diffuse away from the interface through the oxide as has been previously suggested [4, 5], then the predicted peak would again be lower than observed. Based on present calculations it appears that, during hot electron stress, hydrogen atoms desorb from interfacial dangling bonds and are trapped on the silicon side of the interface.

5. Conclusions

We have reported calculations for the relative formation energies of atomic H in bulk Si and SiO$_2$. We have compared theory with experiment in order to develop an understanding of the microscopic behavior of hydrogen during hot electron stress in MOSFETs. Based on the energetics of hydrogen desorption from the interface between silicon and silicon-dioxide, we argue that the hard threshold for hydrogen-related degradation may be considerably lower than the previously assumed value of 3.6 eV. From our analysis, the experiments of Ref. [2] indicate that during hot electron stress Si–H bonds are broken and the atomic hydrogen does not mainly diffuse away into SiO$_2$ but instead is trapped at the interface in a defect configuration with a negative correlation energy.

References

[1] J. W. Lyding, K. Hess, and I. C. Kizilyalli, Appl. Phys. Lett. **68**, 2526 (1996); K. Hess, I. C. Kizilyalli, and J. Lyding, IEEE Trans. Elect. Dev. **45**, 406 (1998).

[2] R. A. B. Devine, J. L. Autran, W. L. Warren, K. L. Vanheusdan, and J. C. Rostaing, Appl. Phys. Lett. **70**, 2999 (1997).

[3] T. Ference, J. Burnham, W. Clark, T. Hook, S. Mittl, K. Watson, and L. Han, IEEE Trans. Elect. Dev. **46**, 747 (1999).

[4] C. Hu, S. Tam, F. Hsu, P. Ko, T. Chan, and K. Terril, IEEE Trans. Elect. Dev. **32**, 375 (1985).

[5] P. M. Lenahan and J. F. Conley Jr., Appl. Phys. Lett. **71**, 3126 (1997).

[6] E. Cartier, D. A. Buchanan, J. H. Stathis, and D. J. DiMaria, J. Non-Cryst. Solids **187**, 244 (1995).

[7] C. G. Van de Walle, P. J. H. Denteneer, Y. Bar-Yam, and S. T. Pantelides, Phys. Rev. **B39**, 10791 (1989); C. G. Van de Walle and W. B. Jackson, Appl. Phys. Lett., **69**, 2441 (1996).

[8] C. G. Van de Walle, Phys. Rev. **B49**, 4579 (1994).

[9] B. Tuttle and C. Van de Walle, Phys. Rev. **B59**, 12884 (1999).

[10] A. Yokozawa and Y. Miyamoto, Phys. Rev. **B55**, 13783 (1997).

[11] P. E. Blochl and J. H. Stathis, Phys. Rev. Lett. **83**, 372 (1999).

[12] B. Tuttle, Energetics and diffusion of hydrogen in SiO_2, Phys. Rev. **B61**, 5429 (2000).

[13] B. Tuttle, Phys. Rev. **B60**, 2631 (1999).

[14] B. R. Tuttle, W. McMahon, and K. Hess, Hydrogen and hot electron creation at the $Si(100)$-SiO_2 interface of metal-oxide-semiconductor field effect transistors, submitted to Supperlatt. Microstruct. **27**, 229 (2000).

[15] P. Hohenburg and W. Kohn, Phys. Rev. **B136**, 869 (1964); W. Kohn and L. J. Sham, Phys. Rev. **A140**, 1133 (1965).

[16] M. Staedele, F. A. Majewski, P. Vogl, and A. Gorling, Phys. Rev. Lett. **79**, 2089 (1997).

[17] N. Johnson, C. Herring, and C. G. Van de Walle, Phys. Rev. Lett. **73**, 1994; *ibid* **74**, 1889 (1995).

[18] F. Himpsel, F. McFeely, A. Taleb-Ibrahimi, J. Yarmoff, and G. Hollimger, Phys. Rev. **B38**, 6084 (1988).

[19] S. E. Rauch, F. J. Guarin, and G. LaRosa, IEEE **EDL-19**, 463 (1998).

[20] M. V. Fischetti and S. E. Laux, IEDM Technical Digest, **305**, (1995), to be published in the *Proceedings of the International Workshop on Silicon Nanoelectronics in Kyoto*, Japan (June, 1999).

[21] K. L. Brower, Phys. Rev. **B42**, 3444 (1990).

[22] K. Hess, B. Tuttle, L. F. Register, and D. K. Ferry, Appl. Phys. Lett. **75**, 3147 (1999).

[23] G. Gerardi, E. Poindexter, P. Caplan, and N. Johnson, Appl. Phys. Lett. **49**, 348 (1986).

Strong isotope effects in the ultraviolet light-induced reactivation of dopants in hydrogenated or deuterated n-GaAs:Si

J. CHEVALLIER, M. BARBÉ

Laboratoire de Physique des Solides et de Cristallogénèse, UMR CNRS 8635, 1 place A.Briand, 92195 Meudon Cedex, France

M. CONSTANT

Laboratoire de Spectrochimie Infrarouge et Raman, UMR CNRS 8516, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq Cedex, France

D. LORIDANT-BERNARD, E. CONSTANT, S. SILVESTRE

Institut d'Electronique et de Microélectronique du Nord, UMR CNRS 8520, BP69, Avenue Poincaré 59652, Villeneuve d'Ascq Cedex, France

The reactivation of silicon donors in hydrogenated n-GaAs:Si under illumination is studied by electrical conductivity and Hall effect experiments at different excitation wavelengths. Ultraviolet illumination at room temperature of hydrogenated n-GaAs:Si by photons with energies above 3.5 eV is found to be an efficient alternative way to reactivate silicon donors. A very weak barrier exists for the dissociation of Si–H(D) complexes under UV excitation. Moreover, a strong isotope effect is observed in the dissociation kinetics of these complexes at 300 K and 100 K for low photon densities. We propose that the UV illumination induces an electronic excitation of Si–H complexes in GaAs. The strong isotope effect is discussed in the light of recent electronic excitation models of Si–H(D) bonds at the surface of silicon.

© 2000 Academic Press

Key words: GaAs, hydrogen diffusion, dopant passivation, dopant reactivation, isotope effects.

1. Introduction

Very often, semiconductors are contaminated by hydrogen during their growth or during some of the processing steps required for the device fabrication. A consequence of this contamination is a passivation of the dopants by hydrogen and one has to find out ways to dissociate the complexes in order to reactivate the dopants. Thermal excitation is the most usual process used but other excitations may enhance the dissociation of H-dopant complexes: minority carrier injection, low-energy electron beam irradiation, hot carrier injection or ultraviolet illumination [1–4]. On the other hand, it might be useful to diffuse hydrogen in certain regions of a device in order to passivate defects or to create low-conductivity patterns by passivation of dopants. In this case, the H-dopant complexes might be thermally, optically or electrically excited during the device

operation conducting to the failure of the device. For this reason, some works have considered whether D-dopant complexes could be more stable than H-dopant complexes under the above excitation processes. Moreover, the investigation of the isotope factor of the H(D)-dopant complex dissociation rate is expected to give some insight into the physical mechanisms responsible for the excitation of the H(D)-dopant complexes. Thermal excitation usually displays a small isotope effect, the D-dopant complexes being slightly more stable than the H-dopant complexes [5]. However, under hot electron injection in reverse bias n-GaAs:Si Schottky diodes, Si–D complexes are significantly more stable than Si–H complexes [3]. A strong isotope effect has been discovered for the stability of Si–H bonds under hot electron excitation at the Si/SiO_2 interface of CMOS transistors [6]. Recently, we found a strong isotope factor in the dissociation kinetics of Si–H and Si–D complexes in GaAs under UV illumination at 300 K [5]. In this work, we shall show that this isotope effect exists at 300 K and also at 100 K.

2. Experimental

For this study, we have used 300 nm thick n^+-GaAs:Si epilayers doped with $n = 3$–4×10^{18} cm^{-3}. They were grown by molecular-beam epitaxy on semi-insulating $\langle 100 \rangle$ GaAs substrates. They have been exposed, at 190 °C, to a rf hydrogen ($t = 6$ h) or deuterium ($t = 8$ h) plasma. The exposure durations were chosen to be long enough to ensure that hydrogen or deuterium diffusion occurs through the whole epitaxial layers of Si-doped GaAs. The sheet carrier concentration and the bulk carrier concentration were deduced from the conventional Hall effect and conductivity experiments. Usually, after hydrogen or deuterium plasma, we remove the first 150 nm of the epilayers by chemical etching in order to suppress the near-surface region which might have experienced some damage due to the plasma exposure. After plasma exposure followed by chemical etching, the carrier concentration is 1–2×10^{17} cm^{-3}. The hydrogenated or deuterated Si-doped GaAs epilayers were then illuminated by using either different harmonic lines of YAG and Ti–sapphire lasers (1 mW cm^{-2}) or the 350 nm emission of an argon laser (light power density in the range 0.8–512 mW cm^{-2}) or the monochromatic light provided by a UV spectrometer (0.5 mW cm^{-2}). For the investigation of the isotope effect, we performed the experiments at two different temperatures: at 300 K and about 100 K. In this case, the sample was mounted on a holder cooled by thermal conduction in a cryostat filled with liquid nitrogen. The sample was illuminated by UV for a given duration and then the Hall effect and the electrical conductivity were measured in the dark, the temperature of 100 K being maintained during all the illumination/electrical measurement cycles.

3. Reactivation of dopants under UV illumination in n-GaAs:Si,H

Figure 1 presents the wavelength dependence of the sheet carrier concentration at 300 K after 3 h of exposure of illumination at a constant light power density of 1 mW cm^{-2}. The important feature of this figure is the existence of a threshold energy of 3.5 eV above which the sheet carrier concentration strongly increases.

In a previous work, we have demonstrated a very good correlation between the increase of the sheet carrier concentration for photon energies above 3.5 eV and the decrease of the absorption intensity related to the Si–H bending mode at 890 cm^{-1} [4]. This definitely proved that the excess of sheet carrier concentration arises from the reactivation of the silicon dopant by dissociation of the Si–H complexes under UV illumination. As expected, in addition to the carrier concentration increase, one also finds a decrease of the electron mobility due to the increase of ionized silicon donor concentration. This is clearly observed in Fig. 2 for the hydrogenated n-GaAs:Si submitted to high incident photon densities. At low photon densities, the electron mobility starts to increase. This might be due to some local diffusion of hydrogen towards scattering centers which are then passivated by hydrogen attachment.
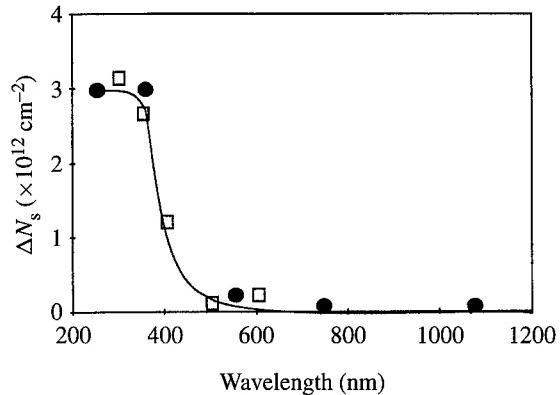
Fig. 1. Illumination wavelength dependence of the extra sheet carrier concentration resulting from the light-induced Si–H complex dissociation in n-GaAs (3 h light exposure): ●, YAG or Ti–sapphire laser illumination; □, monochromatic radiation provided by a UV spectrometer (power density: 1 mW cm^{-2}).
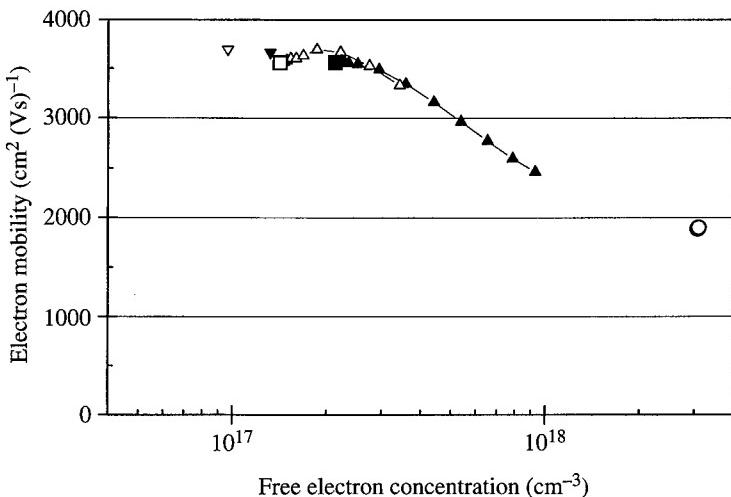
Fig. 2. Evolution of the Hall electron mobility with the Hall carrier concentration resulting from the two parallel conduction paths, the passivated region and the region reactivated by UV illumination: ● (H) and ○ (D), before passivation; ▼ (H) and ▽ (D), after passivation; ■ (H) and □ (D), after passivation and chemical etching; ▲ (H) and △ (D), after passivation and chemical etching followed by UV illumination.

4. Origin of the UV-induced excitation and the strong isotope effects

In Fig. 3, we present the evolution of the extra sheet carrier concentration ΔN_s with the incident photon density for hydrogenated and deuterated n-GaAs:Si at 300 K and 100 K. For densities less than 10^{19} photons cm^{-2}, the illumination duration was kept constant and equal to 2 min while the illumination power density was progressively increased from 0.5 to 32 mW cm^{-2} at 300 K (128 mW cm^{-2} at 100 K). For higher photon densities, we kept constant the highest illumination power density and increased the illumination duration. A power of 512 mW cm^{-2} was used only to reach the highest photon density needed in the case of deuterated n-GaAs:Si at 100 K. We note that, for the hydrogenated and deuterated structures at 300 K, the evolution of ΔN_s does not show any important modification. This means that, in the illumination intensity range that we
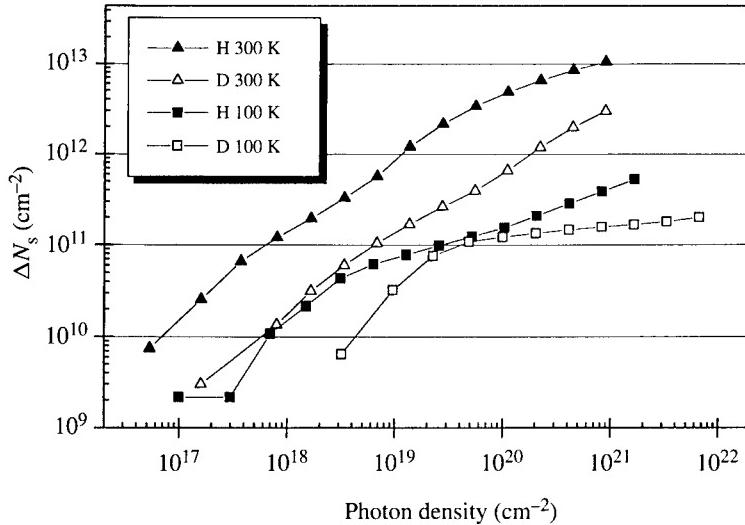
Fig. 3. Photon density dependence of the extra sheet carrier concentration for a hydrogenated and deuterated structure at 300 K and 100 K. The UV illumination comes from the 350 nm emission of an argon laser.

have investigated, the complex dissociation yield is relatively independent of the illumination power density. We shall discuss later the situation of the hydrogenated and deuterated structures at 100 K.

For low photon densities, one observes two major features: a quasilinear variation of ΔN_s with the incident photon density and an important isotope effect in the dissociation yield of Si–H and Si–D complexes at 300 K and 100 K. This result confirms our previous results [5] where the plasma exposed n-GaAs:Si,H epilayers were not etched before UV illumination while those presented here are reported on samples etched on the first 150 nm. This means that the strong isotope factor is related to the Si–H(D) dissociation yield and is not damage-related. At 100 K, the same isotope effect is clearly observed at low photon densities. In addition, we see that the dissociation yield of Si–H complexes is 10 times lower at 100 K than at 300 K. This is a first indication that the barrier energy for dissociation under UV excitation is very small.

At 100 K, the hydrogenated structure exhibits a significant reduction of the dissociation yield for photon densities above 3×10^{18} cm^{-2}. An even more severe effect occurs in the deuterated structure above 2×10^{19} cm^{-2}. We propose to attribute this reduction to a retrapping effect of hydrogen or deuterium on the silicon donors. For low photon densities, all the dissociated Si–H(D) complexes are localized in the near-surface region (100 Å) where the UV illumination is mainly absorbed. However, one can speculate the existence of some surface barrier creating a near-surface barrier region depleted of electrons. After dissociation, the Si–H(D) complexes cannot reform in this near-surface region because of this absence of free electrons. For higher photon densities, Si–H(D) complexes are dissociated deeper in the epilayer in regions where free electrons are available so that retrapping can occur.

The quasi-independence of the dissociation yield of Si–H bonds upon the illumination power density and upon the number of photons for low fluences together with the existence of an important isotope effect independent of the temperature are general characteristics of the UV excitation mechanism. They are somewhat parallel with that of the electronic excitation mechanism of Si–H bonds at the silicon surface when they are excited by scanning tunneling microscopy electrons with energies above the electronic excitation threshold of Si–H bonds (6 eV) [7]. For these reasons, we attribute the UV-induced dissociation of Si–H(D) complexes in GaAs to an electronic excitation of the Si–H bonds. For the physical origin of this electronic excitation, we rule out a mechanism where the minority carriers created by UV illumination would be directly responsible

for the enhanced dissociation of Si–H complexes because the threshold energy of 3.5 eV of the complex dissociation is well above the bandgap energy of GaAs and is different from the onset of its strong absorption at 3.1 eV. Electrons resulting from the UV illumination have an energy of at most 2 eV before their thermalization. This energy is probably not high enough to provide an electronic excitation of the Si–H complexes contrary to the case of impact dissociation by hot electrons either injected in highly reverse bias Schottky n-GaAs:Si,H diodes [3] or created by electron-beam exposure [8]. We propose that the electronic excitation of Si–H complexes arises from a UV-induced photodissociation process where the Si–H bonds are excited to a quasi-antibonding state. Experimentally, it has been established that photochemical hydrogen desorption occurs at the surface of Si(100) and Si(111) under UV illumination with photons of 7.9 eV energy indicating a direct electronic excitation of Si–H bonds at the silicon surface by UV illumination [9, 10]. Very recently, it has been found by *ab initio* molecular dynamics calculations that UV photons of 4 eV are able to give rise to an electronic excited state of Si–H complexes in GaAs [11].

An easy explanation of the isotope effect would be the smaller zero-point energy of Si–D bonds compared with Si–H bonds. We rule out this explanation because we would expect an increase of the isotope factor as the temperature decreases which is not the case. The strong isotope effect we observe can be understood in the frame of Menzel and Gomer's model [12] used to explain the strong isotope effect in the stability of Si–H(D) bonds at the surface of silicon illuminated by photons of 7.9 eV [10] or electronically excited by hot electrons [7]. Due to its stronger inertia, deuterium will move more slowly than hydrogen after UV excitation and its electronic excited state will be more heavily quenched. Consequently, deuterium will have a smaller probability of dissociation compared with hydrogen.

5. Conclusion

In conclusion, we have shown that the dissociation of Si–H(D) complexes in GaAs under UV illumination can be explained by an electronic excitation process with a UV-induced photodissociation. A strong isotope effect is observed in the dissociation kinetics of Si–H and Si–D complexes under UV excitation at 300 K and also at 100 K provided the retrapping of hydrogen on silicon donors does not play a significant role. This isotope effect can be explained according to the models presently used to explain H/D isotope effect in hot electron and UV photon-induced hydrogen desorption at the surface of silicon. A more detailed study of the temperature dependence of the dissociation yield is necessary to provide the barrier energy for dissociation under UV excitation.

References

[1] N. M. Johnson, C. Doland, F. Ponce, J. Walker, and G. Anderson, Physica **B170**, 3 (1991).

[2] H. Amano, M. Kito, K. Hiramatsu, and I. Akasaki, Jpn. J. Appl. Phys. **28**, L2112 (1989).

[3] E. Constant, D. Loridant-Bernard, S. Mézière, M. Constant, and J. Chevallier, J. Appl. Phys. **85**, 6526 (1999).

[4] D. Loridant-Bernard, S. Mézière, M. Constant, N. Dupuy, B. Sombret, and J. Chevallier, Appl. Phys. Lett. **73**, 644 (1998).

[5] J. Chevallier, M. Barbé, E. Constant, D. Loridant-Bernard, and M. Constant, Appl. Phys. Lett. **75**, 112 (1999).

[6] J. W. Lyding, K. Hess, and I. C. Kizilyalli, Appl. Phys. Lett. **68**, 2526 (1996).

[7] Ph. Avouris, R. E. Walkup, A. R. Rossi, T.-C. Shen, G. C. Abeln, J. R. Tucker, and J. W. Lyding, Chem. Phys. Lett. **257**, 148 (1996).

[8] E. Constant, S. Silvestre, D. Bernard-Loridant, M. Constant, and J. Chevallier, Superlatt. Microstruct. **27**, 431 (2000).

[9] A. Pusel, U. Wetterauer, and P. Hess, Phys. Rev. Lett. **81**, 645 (1998).

[10] T. Vondrak and X.-Y. Zhu, Phys. Rev. Lett. **82**, 1967 (1999).

[11] Y. Miyamoto, O. Sugino, and Y. Mochizuki, Appl. Phys. Lett. **75**, 2915 (1999).

[12] D. Menzel and R. Gomer, J. Chem. Phys. **40**, 1164 (1964).

Tunneling spectroscopy of voltage tunable quantum wires

G. PLONER, E. GORNIK

*Institut für Festkörperelektronik, Mikrostrukturzentrum, Technische Universität Wien, Floragasse 7,
A-1040 Vienna, Austria*

Resonant tunneling spectroscopy is used to investigate the tuning range for the one-dimensional subband spacing of side-gated quantum wires. We introduce a simplified selective depletion scheme for the implementation of a resonant tunneling device. From the analysis of the differential tunneling conductance obtained for a single-wire device we conclude that the energetic spacing for the one-dimensional subbands can be varied from effectively 0 to about 6 meV. Measurements in magnetic fields directed parallel and perpendicular to the tunnel current confirm the one-dimensional nature of the tunneling processes as well as the order of magnitude of the subband spacing by comparison of the tunneling characteristics with a model calculation that assumes a parabolic confinement.

© 2000 Academic Press

Key words: quantum wire, magnetotunneling spectroscopy.

Introduction

An interesting possibility for the practical application of quantum wires is their use for detecting devices in the THz range. This is because the energetic spacing of their one-dimensional (1D) subbands is typically in the 1–10 meV range and can be tuned with a gate by varying the strength of the lateral confinement. The present study of single quantum wire devices is motivated by the question: which tuning range can be achieved for the 1D subband energies by using a combination of deep mesa etching and side gates for the variation of the lateral confinement? We chose tunneling spectroscopy as the method to investigate the properties of side-gated wires, because it is the most sensitive technique for the study of the spectral properties of low-dimensional electron systems [1]. The main technological difficulty encountered in the fabrication of resonant tunneling devices concerns the Ohmic contacts. They have to be established independently to two electron channels, which are usually separated by a tunneling barrier with a thickness of several hundred Å. Several successful methods for the fabrication of selective contacts have been proposed. Two of them are restricted to special tunneling geometries and rely either on the formation of shallow Ohmic contacts [1], or on a series of intricate etching steps leading to the formation of free-standing GaAs bridges [2, 3]. Another very successful fabrication scheme uses patterned back-gates, implemented either by thinning the wafer to a few micrometers [4, 5] or by *in situ* ion implantation followed by overgrowing the pre-processed back-gate pattern with the active double-layer electron system [6, 7]. There are also some recent developments in device technology suggesting that resonant tunneling transistors might also have promising applications for room-temperature operation in ultra-large-scale integrated circuits [8, 9]. This indicates that it is desirable to find a way to simplify the usual procedures for selective contact formation without the need of a two-sided patterning of the samples or intermediate implantation steps. As an attempt to achieve this we used a
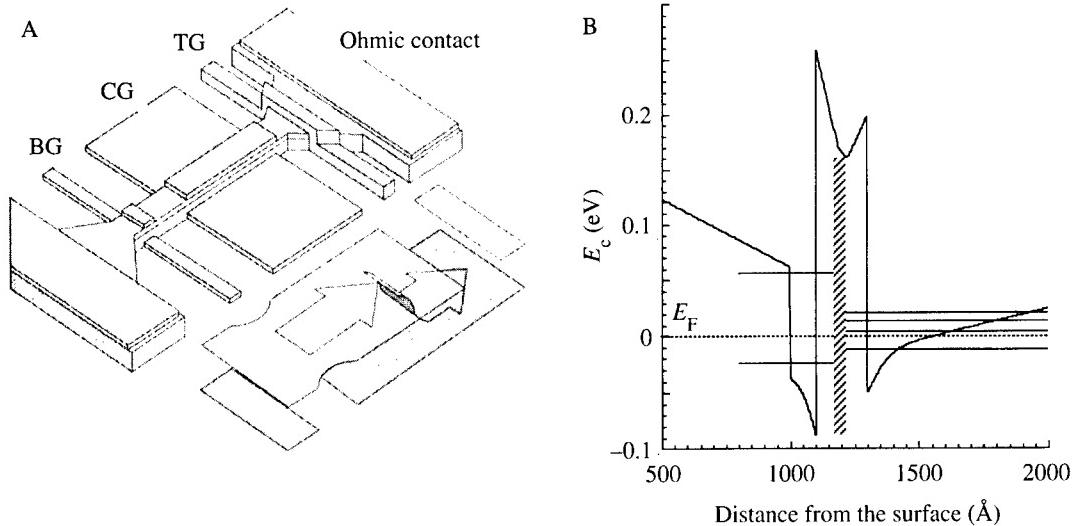
Fig. 1. A, Schematic picture of the device layout (not to scale). The lower part of the figure symbolizes the current path for reverse bias in a situation where the top (TG) and bottom (BG) gates have pinched off the electron gases at their respective positions. B, Conduction band profile and self-consistently calculated energy levels of the tunneling structure used to demonstrate the present device principle. The shaded region indicates the δ-doping of the barrier.

modified device scheme, which is both simple and flexible and uses self-aligned side gates for the selective depletion of tunnel-coupled electron gases. In this report we present 1D–1D tunneling data obtained with a single-quantum-wire tunneling device, which has been fabricated with this technique. The same fabrication method is shown to be suitable for imposing a tunable lateral confinement of the active region for the tunneling device.

Sample layout

Figure 1A shows a two-terminal device designed for testing the side-gate selective depletion scheme. The active area of the device consists of an approximately 1 μm wide channel, which extends between two large Ohmic contacts established to both electron gases simultaneously. The upper electron channel is separated from one of the contacts by a conventional top Schottky gate on the one end of the active region (in the following referred to as the top gate and labeled TG in Fig. 1A). At the other end of the device a side gate is positioned in close proximity to the walls of the etched channel (bottom gate, BG in the figure). When the double electron layer structure is suitably chosen, the bottom gate will locally deplete the lower of the two electron gases, while the upper one still remains conducting. This leads to independent contacts to the two electron channels as depicted in the lower part of Fig. 1A. The bottom gate is fabricated using a self-alignment technique that is described in detail in Ref. [10]. The same technique can be used to align an additional central gate (CG in the figure) intended to impose a tunable lateral confinement on the active tunneling region. The main advantage of the described device scheme consists of the simplicity of the fabrication process, which is entirely planar and relies on standard optical lithography and wet chemical etching.

In the following we present some results for a single-barrier tunneling structure. The sample structure consists of a strained 100 Å wide $In_{0.15}Ga_{0.75}As$ quantum well (referred to also as the 'upper' channel in the following), followed by a δ-doped 200 Å thick $Al_{0.3}Ga_{0.7}As$ barrier and a 7500 Å thick, nominally undoped GaAs buffer. The triangular well formed at the AlGaAs–GaAs junction will be called the 'lower' channel in the following. The electron concentrations in the upper and lower 2DEGs were determined by Shubnikov–de

Haas measurements to be 6.2×10^{11} cm^{-2} and 3.1×10^{11} cm^{-2}, respectively. The band profile and energy levels, calculated self-consistently using the experimentally found parameters, are shown in Fig. 1B.

The experimental situation is depicted in the lower part of Fig. 1A. A negative voltage is applied to the top and bottom gates such that the two electron gases (symbolized by the two parallel planes) are pinched off at the positions indicated in the drawing. The following results are obtained by the technique of non-equilibrium tunneling spectroscopy [1], i.e. a bias is applied between the two separately contracted electron channels and the corresponding differential tunneling conductance (DTC) is measured using a lock-in technique.

1D–1D tunneling

Figure 2 shows a set of differential conductance traces dI/dV_b for various values of the central gate voltage V_{CG}. The data were obtained for a device whose electron channel had a lithographic width of 1 μm and a length of 20 μm. For $V_{CG} = 0$ V (inset) the peaks observed on the negative bias side of the dI/dV_b curve can be attributed to resonant tunneling processes starting from the lowest occupied 2D subbands of the InGaAs well and ending in one of the 2D subbands of the GaAs–AlGaAs junction [10]. The broad peak at positive V_b corresponds to tunneling out of the lowest subband of the junction into the first excited subband of the InGaAs well. Note that there is also a slight shift of several millivolts of the entire tunnel characteristics towards negative bias. This is due to the two-terminal structure of our test device and results from the series resistance induced by the bottom gate.

As the central gate voltage becomes increasingly negative, a wealth of additional structure appears in the differential conductance curves of Fig. 2. This can be attributed to resonant tunneling between the 1D subbands that develop when the electron channels become more and more confined by the central gate. For a proper interpretation of the 1D–1D tunneling structure it is necessary to consider the fact that tunneling occurs both in a region influenced by the central gate and another region unaffected by this gate. These different areas of the device are also schematically indicated in the left inset of Fig. 2. In the gated region the lower channel not only gets increasingly confined but also will be partially depleted as V_{CG} becomes more negative. The first pronounced peak of the dI/dV_b traces around $V_b = -0.6$ V is seen to remain mainly unchanged with varying V_{CG}. One may therefore attribute this feature to tunneling in the ungated (2D) regions of the device. The broad fine-structured features in the differential conductance traces result from tunneling between the 1D states of the gated region. Due to the depletion of the lower channel, the corresponding resonant structure is shifted towards more negative bias voltage. The gradual depletion of the lower channel also becomes obvious from the suppression of the peak in forward direction (on the positive bias side of the dI/dV_b traces).

The fine structure of the dI/dV_b can be analyzed in terms of the model of Mori *et al.* [11]. It is based on the transfer Hamiltonian formalism and describes 1D–1D tunneling processes between laterally confined states, assuming parabolic confinement for the emitter and the collector states. In order to apply the main result of this model to our data it is necessary to account for the fact that the applied bias voltage V_b is not equal to the voltage drop across the tunneling barrier (denoted by V_b^* in the following). This is mainly due to the already mentioned series resistance induced in the current path by the bottom gate (Fig. 1A). The first peak in dI/dV_b around -0.6 V, previously assigned to 2D–2D tunneling in the ungated regions of the device, turns out to be a useful means to determine the voltage drop across the barrier and thus to find the correct relationship between V_b and V_b^*. In order to do this a magnetic field is applied in the direction of the tunneling current (perpendicular to the sample). The splitting of the 2D subbands into Landau levels (LLs) of a definite energy spacing results in a characteristic splitting of the tunneling peak under consideration, which should be largely independent of the central gate voltage. This is shown in Fig. 3A, which displays two typical tunneling characteristics for $V_{CG} = 0$ V and $V_{CG} = -0.8$ V in a magnetic field of 5 T. In this figure the first 2D–2D peak of the dI/dV_b curve is seen to split into a series of peaks due to inter LL tunneling. The overall structure of the LL resonance is seen to be indeed independent of V_{CG} (besides
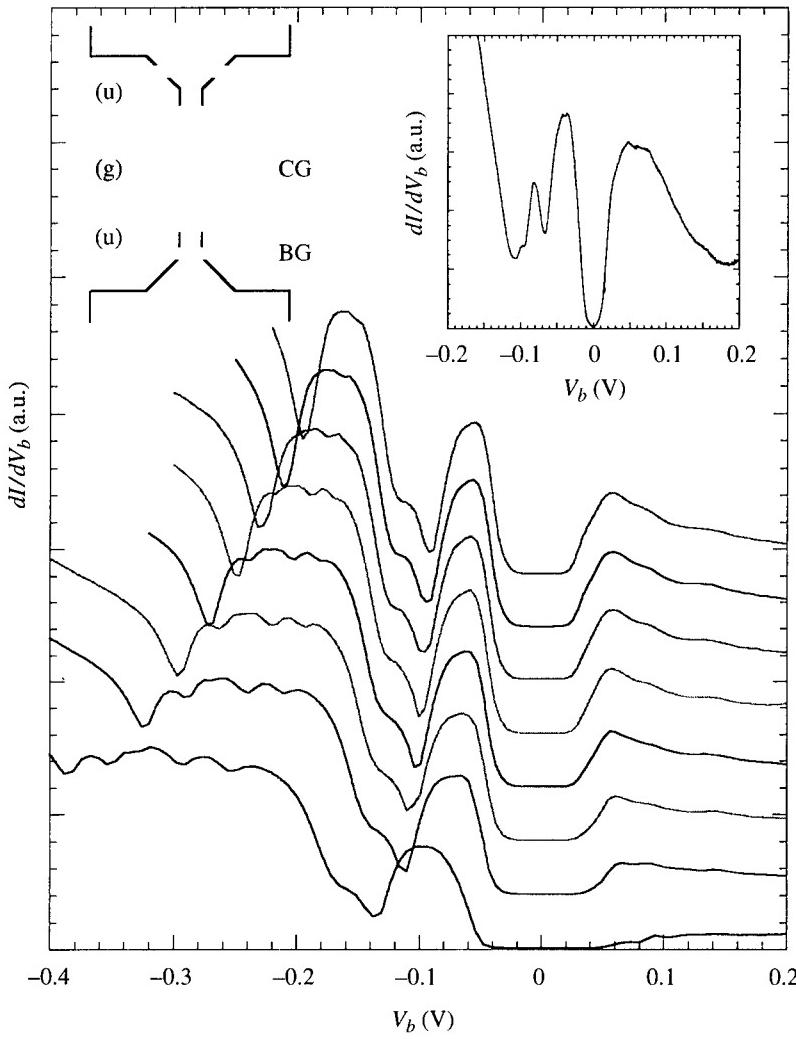
Fig. 2. Differential conductance dI/dV_b, measured at $T = 1.9$ K for various values of V_{CG} (from top to bottom: $V_{CG} = -0.6, -0.65, -0.7, -0.75, -0.8, -0.85, -0.9, -1.0$ V). Curves are vertically offset. The right inset shows the differential conductance dI/dV_b for the 2D–2D case at $V_{CG} = 0$ V. The left inset symbolizes the top view of the device. Gated (g) and ungated (u) regions of the device lead to different parts of the tunneling spectra as described in the text.

slight differences in the relative heights of the various sub-peaks, which can be attributed to different non-resonant background currents). A relationship between the measured peak spacing and the LL energy has to be obtained by calculating the expected tunneling spectrum. In the calculation a Gauss curve is assigned to each tunneling transition between two LLs according to the results given in Ref. [12]. Each resonance occurs at a certain voltage drop V_b^* across the barrier, where the corresponding LLs are energetically aligned. After adding the various Gaussians, the resulting $I(V_b^*)$ curve is superimposed onto some continuous background and finally differentiated with respect to V_b^*. An example of a calculated differential tunneling characteristic is shown in Fig. 3B. Since we are mainly interested in the peak distance of the most prominent features of the DTC, the calculation does not take into account the voltage shift of the experimental tunneling traces and neglects those tunneling transitions ending up in the higher LLs of the second subband of the GaAs–AlGaAs
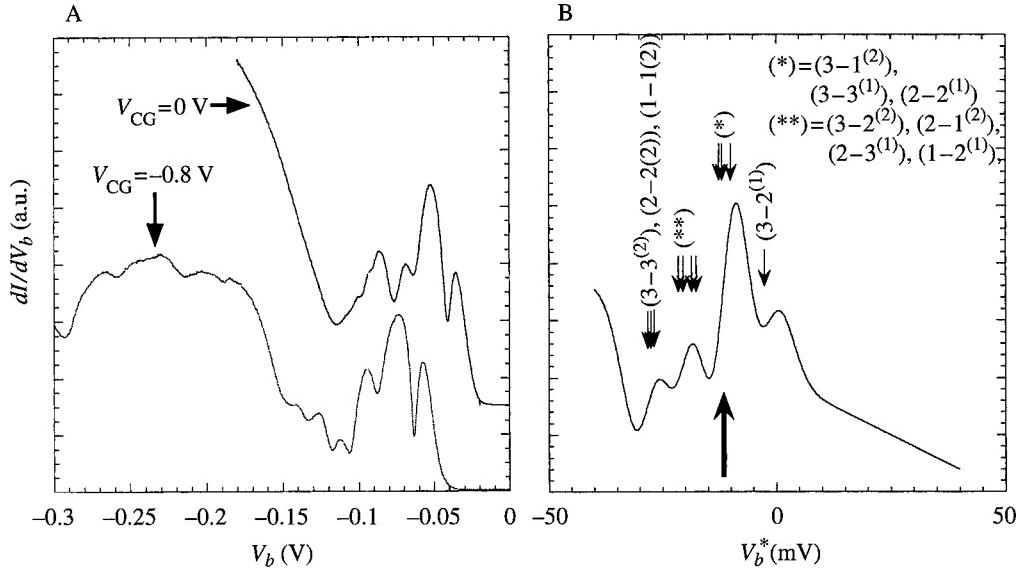
Fig. 3. A, Differential conductance corresponding to $V_{CG} = 0$ V and -0.8 V, recorded in a magnetic field parallel to the tunneling current with field strength of 5 T. B, Calculated DTC for a magnetic field of 5 T. Downward arrows indicate the various transitions between Landau levels that have been taken into account in the calculation. The notation $i - j^{(k)}$ means a transition from the ith LL in the InGaAs well into the jth LL of the kth subband of the GaAs–AlGaAs junction. The upward arrow indicates the voltage position of the first 2D–2D resonance without magnetic field. Note that here dI/dV_{b}^{*} is plotted against the voltage drop across the tunneling barrier V_{b}^{*}.

junction. It is also assumed that the non-resonant background current depends quadratically on V_{b}^{*}, which is only a rough approximation to the true behaviour. In spite of these simplifications, the calculated DTC is seen to reproduce the measured features of the ungated peak in Fig. 3A quite well. A comparison of the calculated and measured peak spacing allows us to obtain a scaling factor for the bias voltage axis for each value of the central gate voltage. In other words, this scaling procedure transforms the measured $I(V_{b})$ into $I(V_{b}^{*})$.

Another point to be noted here is that the observed resonant structures are rather broad. For example, in the calculated DTC of Fig. 3B a LL width (FWHM) of 5 meV had to be assumed in order to reproduce the experimental curves. The observed large linewidth is mainly due to the heavily doped tunneling barrier, which causes elastic scattering of tunneling electrons and also a low mobility in the two electron channels. The presence of elastic scattering, on the other hand, leads to the inter-LL tunneling peaks used to re-scale the V_{b}-axis. The corresponding tunneling transitions would be forbidden if energy and momentum were exactly conserved during a resonant tunneling process [12].

Finally we note that the energetic spacing of the 1D subbands is expected to be different in the InGaAs well and the GaAs–AlGaAs heterojunction. Due to the geometry of the sample layout and the different carrier densities, the lower channel will become more strongly confined than the upper one. This agrees with a simple estimate of the confinement potential based on electrostatic arguments [13].

It has been demonstrated [2, 11] that in this situation the tunneling spectrum $I(V_{b})$ consists of a series of resonant peaks, which are grouped to 'compound' peaks containing all 1D–1D tunneling transitions between initial and final 1D states of the same parity. The spacing between 'compound' peaks reflects the subband spacing of the more strongly confined electron channel. Superimposed on the 'compound' peaks there is some fine structure due to the single-resonant transitions contributing to the compound resonance. The spac-
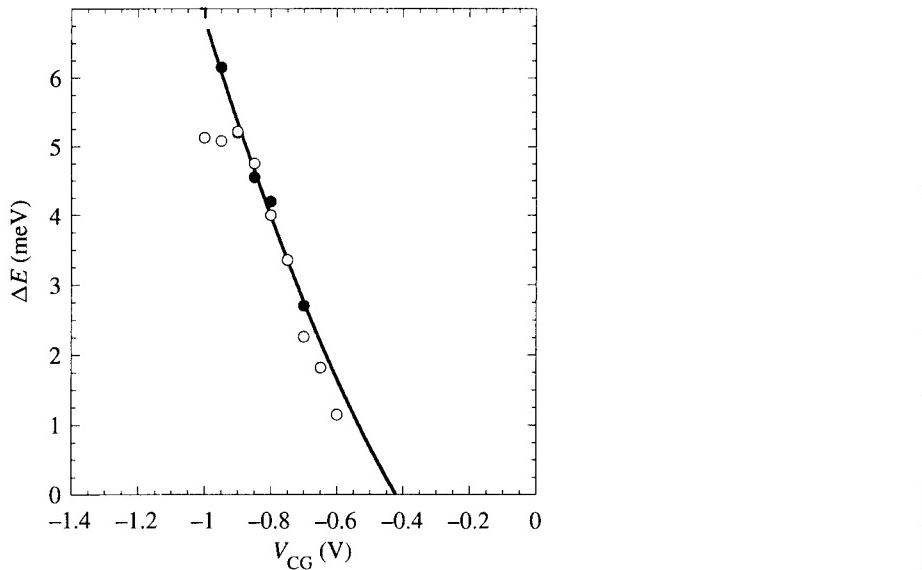
Fig. 4. Subband spacing of the lower channel as a function of the central gate voltage. Solid and open symbols were obtained from different pairs of peaks of the 1D–1D tunneling characteristics shown in Fig. 2. The solid line is a quadratic fit to the data and serves as a guideline to the eye.

ing of these fine-structured features reflects the subband energy of the weakly confined channel. As discussed in Ref. [10], it is these 'compound peaks' that are observed in the dI/dV_b characteristics of Fig. 2. After the V_b-axis has been rescaled, the voltage difference between two peaks then approximately equals twice the subband spacing of the lower GaAs–AlGaAs junction. The superimposed fine structure due to the smaller subband spacing of the upper InGaAs channel cannot be resolved in the present experiment due to the large intrinsic width of the single resonances.

Figure 4 shows the derived subband spacing of the lower channel as a function of central gate voltage. Solid and open symbols were determined from two different pairs of peaks of the 1D–1D tunneling features in Fig. 2. For $V_{CG} \geq -1$ V one enters the pinch-off regime of the lower channel. In this regime the relation between the measured V_b and V_b^* is no longer linear and the outlined analysis scheme becomes invalid. This also becomes evident in the $V_{CG} = -1$ V curve in Fig 2. Increasing V_{CG} beyond -0.9 V is seen to have the only effect of stretching the tunneling characteristics over a larger bias voltage range without revealing any few features. This is the reason why the subband energies seem to saturate at the highest values of V_{CG} in Fig. 4.

The results show that with the present device it is possible to achieve quite large subband energies (in the lower channel), ranging between 0 and 6 meV.

The estimated magnitude of the 1D subband spacing is also confirmed by temperature-dependent measurements. An example is shown in Fig. 5. Here the DTC is plotted for a different device with a slightly wider channel (1.2 μm) for a number of temperatures. Again one observes the 'ungated' 2D–2D tunneling peak followed by the structure corresponding to 1D–1D tunneling in the gated region of the device. The 1D–1D peaks are seen to be largely smeared out at about 50 K, i.e. when the thermal energy is of the same order of magnitude as the 1D subband energy. Since $k_B T \approx 4.3$ meV at 50 K, the subband energy is seen to be of the same magnitude even for a 1.2 μm wide device.
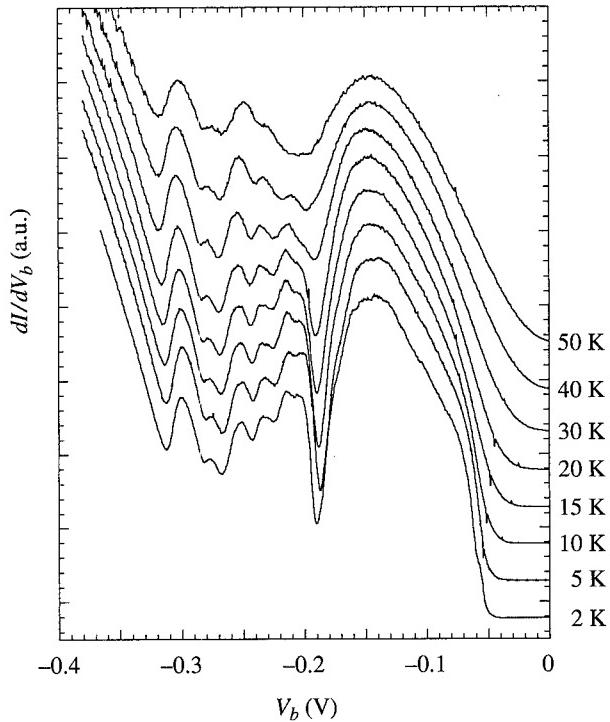
Fig. 5. DTC measured at various temperatures for a device with a 1.2 μm wide channel. The central gate voltage in this case was $V_{CG} = -2.6$ V; pinch off occurs for this device beyond $V_{CG} = -3.5$ V.

Measurements in nonzero magnetic fields

The following gives a brief account of the results obtained when a magnetic field is applied parallel or perpendicular to the direction of the tunneling current. Both types of experiment give additional evidence that the structure observed in the dI/dV_b curves, when the central gate is negatively biased, in indeed due to tunneling between 1D states.

Figure 6 shows how the peak positions of the 1D–1D tunneling structure in Fig. 2 evolve with a magnetic field applied parallel to the tunnel current. The central gate voltage in this case was $V_{CG} = -0.8$ V. An example of typical dI/dV_b trace at $B = 5$ T is shown in Fig. 3A. As can be seen in this figure, there is no significant change in the qualitative features of the respective tunneling characteristics. Figure 6 shows moreover that even in the peak positions no significant change occurs as long as the magnetic field strength is smaller than 3 T. This is exactly what one expects, if the tunneling indeed occurs between 1D levels. A magnetic field perpendicular to the sample surface leads to a B-dependent subband spacing according to $\hbar\omega = \sqrt{(\hbar\omega_0)^2 + (\hbar\omega_c)^2}$, where ω_c denotes the cyclotron frequency and ω_0 is the spacing of the 1D subbands without magnetic field. At low magnetic fields the electrostatic confinement (characterized by ω_0) consequently dominates the additional confinement introduced by the magnetic field, whereas at high B-values the opposite is true. As already discussed by Mori *et al.* [11, 14], in the first case tunneling occurs between 1D states of equal parity, in the second case mainly between magnetoelectric hybrid (Landau) levels of identical LL index. There should therefore occur a crossover between the two cases when the magnetic confinement begins to dominate over the electrostatic one. One expects additional peaks beyond this crossover, since tunneling is now also allowed between magnetoelectric hybrid levels of the same index. As seen in Fig. 6, there is indeed some evidence of such a crossover between 2.5 and 3.5 T. At lower magnetic fields the peak po-
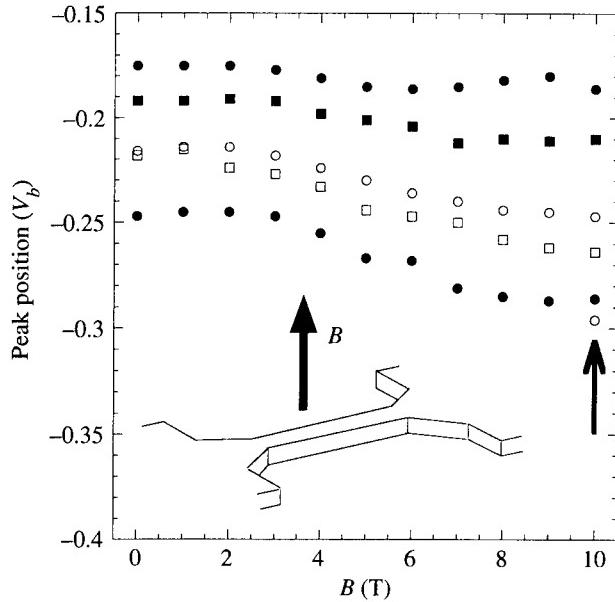
Fig. 6. Peak positions of the 1D–1D resonant structure in the differential conductance obtained for the device with 1 μm channel width and plotted versus magnetic field strength. The orientation of the magnetic field with respect to the device geometry is drawn in the inset. The peak positions are directly obtained from the dI/dV_b traces without re-scaling the bias voltage axis.

sitions are nearly unchanged. For fields $\geq$ 4 T the peak distance is seen to increase. However, no additional structure can be observed except for the 10 T trace where a small peak develops at the outermost boundary of the 1D tunneling structure (marked by an arrow in Fig. 6). This means that here, again, the resolution of the finer features of the tunneling spectra is impossible due to the large linewidth of the single tunneling transitions. The position of the crossover regime in Fig. 6 lies between 2.5 and 3.5 T. The cyclotron energy varies between 4.25 and 5.95 meV in this magnetic field interval, which is again in agreement with the earlier found 1D level spacing of about 4 meV at $V_{CG} = -0.8$ V (cf. Fig. 4). This is another indication that the previous estimates of the subband energies give the correct order of magnitude.

We now briefly discuss the DTC obtained when the magnetic field is oriented perpendicular to the tunneling direction. According to the transfer Hamiltonian model of Mori *et al.* [11, 14], there should be a characteristic difference between the DTC recorded with the magnetic field directed along the wire axis or directed perpendicular to (but in the same plane as) the wire. Figure 7 shows a sequence of differential conductance traces, obtained at $T = 2$ K in a magnetic field directed as indicated in the schematic drawing of the device. Besides some additional broadening of the main resonances, one also observes a shift of the various peaks towards higher bias voltage, accompanied by a variation in the relative peak heights. Thereby the 'center of mass' of the entire tunneling structure seems to be shifted to higher V_b. Again, this is in principal agreement with the general theoretical investigation of 1D–1D tunneling given by Mori *et al.* [11]. If the direction of B is along the wire, one expects to observe additional structures in the 1D–1D DTC due to the single 1D–1D transitions, as well as a change in their relative amplitudes with increasing B. The corresponding measurement, however, does not show a similar fine structure for our samples, but rather gives almost identical curves as in Fig. 7. As above for the case of B parallel to the tunneling direction, this can be attributed to the large linewidth of the tunneling resonances which does not allow us to resolve the single transitions contributing to each main peak in dI/dV_b. Therefore, both magnetic field directions only reflect
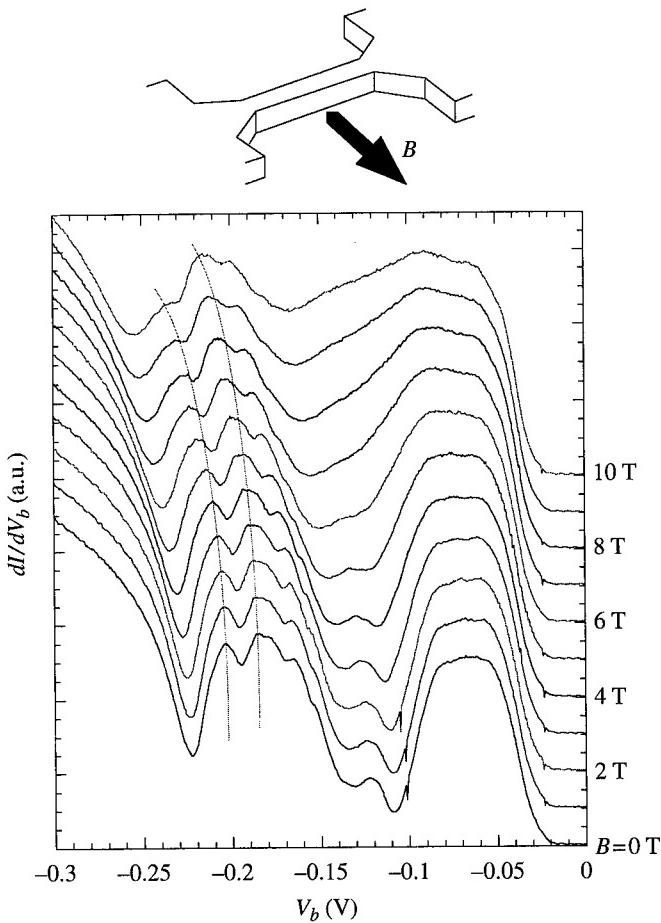
Fig. 7. dI/dV_b traces in a magnetic field perpendicular to the tunneling current, measured for the device with 1 μm channel width. The direction of B is indicated in the schematic picture on top of graph.

the behaviour of the gross structure of the tunneling characteristics, which according to the model of Mori *et al.* [11] should indeed give almost identical results for both magnetic field directions.

Summary

The above considerations demonstrate that the side-gate technique indeed gives a simple and flexible possibility to implement both a selective depletion scheme and a lateral confinement which is very effectively tuned by the voltage applied to the central gate. The tuning range of the 1D sublevel spacing has been found to lie between effectively 0 and about 6 meV, allowing the study of 2D–2D and 1D–1D tunneling processes on one single device. The above selective depletion scheme can be applied to a variety of mesa geometries and is not restricted to two-terminal configurations. The only essential prerequite for the introduced side-gate technique is to have sufficiently different carrier densities in the two barrier-separated electron gases. Our results also show that the use of gate-tunable etched quantum wire structures allows us to obtain 1D subband energies, which are indeed in a range interesting for the application of quantum wires as detecting devices

in the THz domain. An investigation of the photoconductive properties of gate-tunable single or multiple quantum channel devices is the subject of ongoing investigations.

References

[1] J. Smoliner, W. Demmerle, E. Gornik, and G. Weimann, Phys. Rev. Lett. **63**, 2116 (1989).

[2] J. Wang, P. H. Beton, N. Mori, L. Eaves, H. Buhmann, L. Mansouri, P. C. Main, T. J. Foster, and M. Henini, Phys. Rev. Lett. **73**, 1146 (1994).

[3] P. H. Beton, J. Wang, N. Mori, L. Eaves, P. C. Main, T. J. Foster, and M. Henini, Phys. Rev. Lett. **75**, 1996 (1995).

[4] J. P. Eisentein, L. N. Pfeiffer, and K. W. West, Appl. Phys. Lett. **57**, 2324 (1990).

[5] S. Q. Murphy, J. P. Eisenstein, and L. N. Pfeiffer, Phys. Rev. **B52**, R14825 (1995).

[6] K. M. Brown, E. H. Linfield, D. A. Ritchie, G.A C. Jones, M. P. Grimshaw, and M. Pepper, Appl. Phys. Lett. **64**, 1827 (1994).

[7] B. Kardinal, C. H. W. Barnes, E. H. Linfield, D. A. Ritchie, K. M. Brown, G. A. C. Jones, and M. Pepper, Phys. Rev. Lett. **76**, 3802 (1996).

[8] J. A. Simmons, M. A. Blount, J. S. Moon, S. K. Lyo, W. E. Baca, J. R. Wendt, J. L. Reno, and M. J. Hafich, J. Appl. Phys. **84**, 5626 (1998).

[9] J. S. Moon, J. A. Simmons, M. A. Blount, J. L. Reno, and M. J. Hafich, Appl. Phys. Lett. **74**, 314 (1999).

[10] G. Ploner, H. Hirner, T. Maier, G. Strasser, and E. Gornik, Appl. Phys. Lett. **74**, 1758 (1999).

[11] N. Mori, P. H. Beton, J. Wang, and L. Eaves, Phys. Rev. **B51**, 1735 (1995).

[12] W. Demmerle, J. Smoliner, G. Berthold, G. Weimann, and W. Schlapp, Phys. Rev. **B44**, 3090 (1991).

[13] G. Ploner, H. Hirner, T. Maier, G. Strasser, and E. Gornik, in *Proceedings of the 13th International Conference on the Electronic Properties of Two-dimensional Electron Systems 1st–7th August 1999, Ottawa, Canada*, to be published in Physica E.

[14] N. Mori, P. H. Beton, J. Wang, and L. Eaves, Phys. Rev. **B52**, 1504 (1995).

Magnetic anticrossing of 1D subbands in ballistic double quantum wires

M. A. Blount, J. A. Simmons[†]

Sandia National Laboratories, Albuquerque, New Mexico, 87185, U.S.A.
Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87131, U.S.A.

J. S. Moon, S. K. Lyo, J. R. Wendt, J. L. Reno

Sandia National Laboratories, Albuquerque, New Mexico, 87185, U.S.A.

We study the low-temperature in-plane magnetoconductance of vertically coupled double quantum wires. Using a novel flip-chip technique, the wires are defined by two pairs of mutually aligned split gates on opposite sides of a ≤ 1 micron thick AlGaAs/GaAs double quantum well heterostructure. We observe quantized conductance steps due to each quantum well and demonstrate independent control of each 1D wire. A broad dip in the magnetoconductance at ~6 T is observed when a magnetic field is applied perpendicular to both the current and growth directions. This conductance dip is observed only when 1D subbands are populated in both the top and bottom constrictions. This data is consistent with a counting model whereby the number of subbands crossing the Fermi level changes with field due to the formation of an anticrossing in each pair of 1D subbands.

© 2000 Academic Press

Key words: quantum point contact, magnetoconductance, tunneling, anticrossing.

Coupled low-dimensional electronic systems have become of increasing interest recently due to advances in fabrication technology. While much attention has been focused on closely coupled two-dimensional electron systems and coupled quantum dots, little work has been done on coupled quantum wires. Early work focused on tunneling between wires defined laterally by planar lithography, in which case the lateral tunneling is weak and the barrier profile poorly defined [1]. More recently, vertically coupled quantum wires (VCQWs) were defined by mesa-etching a double quantum well heterostructure [2]. While this last approach enables a well-defined tunnel barrier, control over the 1D wire widths is limited, and in any event cannot be performed independently for the two wires. Finally, VCQWs have also been fabricated using a split-gate, with an additional mid-gate within the split-gate opening [3]. The split gate defines the coupled quantum wires, while the midgate controls the relative electron density between the two wires.

In this work, portions of which have appeared previously [4], we study a pair of vertically coupled quantum point contacts (VCQPCs) in a closely coupled GaAs/$Al_{0.3}Ga_{0.7}As$ double quantum well (DQW). The DQW
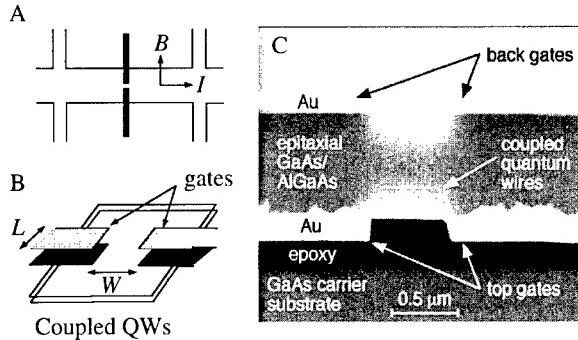
Fig. 1. A, Schematic of experimental geometry for magnetoconductance measurement. B, Split gate geometry. C, Scanning electron micrograph of a test structure. The actual sample was only 3000 Å thick, so that the front and back split gates were of equal distance from the electron layers.

had 200 Å wells separated by an 11 Å barrier. The two occupied subbands of the DQW had densities of 2.5 and 1.2×10^{11} cm^{-2}, and had a total mobility of 5.8×10^5 cm^2 V^{-1} s^{-1}, yielding a mean free path of ~ 10 μm. A symmetric–antisymmetric energy gap of $\Delta_{SAS} = 1.7$ meV was measured. We used the epoxy-bond-and-stop-etch (EBASE) technique, described elsewhere [5], to define a pair of vertically aligned split gates by electron beam lithography *on each side* of the DQW heterostructure, whose thickness is only 3000 Å. This process allows sub-tenth-micron alignment of top and back gates, enabling independent control of the two quantum wires, and has no deleterious effect on sample mobility. In the finished sample both the top and back split gates are placed only $\sim$1500 Å from the DQW. Figure 1A and B show schematics of the sample, and Fig. 1C shows a cross-sectional scanning electron micrograph of a dual split-gate test structure of similar geometry, but having a larger thickness of ~ 1 μm. The width of our split-gate opening was 0.5 micron, and its length was nominally 0.05 microns.

By controlling the relative values of top and back split-gate voltages V_T and V_B, the sample can be tuned into five different regimes depending on the widths of the individual quantum wires. These regimes are: (i) 2D–2D; (ii) 2D–1D; (iii) 1D–1D; (iv) 1D–depleted; and (v) both depleted. Figure 2 shows the sample resistance as a function of V_T, with $V_B = 0$. A number of features appear which correspond to transitions between the different regimes. At $V_T = 0$, both channels are 2D. As V_T is made increasingly negative, the channel resistance shows a clear shoulder at -0.3 V followed by a sharp increase at -0.7 V. These two features are due to the sequential depletion of first the top and then the bottom QW, in the regions *directly beneath* the top split gate. At $V_T \approx -0.7$ V, where electrons beneath the top split gate become completely depleted from *both* QWs, coupled 1D channels are formed in both electron layers. This marks the transition from regime (ii), the 1D–2D case, to regime (iii), the 1D–1D case.

As V_T is swept further, the QPC widths narrow. Due to the fact that the split gate separation is comparable to the electron layer depth, the top wire narrows more rapidly than the bottom wire. At $V_T = -2.6$ V, the top wire completely pinches off, leaving only the bottom wire. This corresponds to the transition between regimes (iii) and (iv), and is marked by weak plateau and a change in the slope of the resistance. As V_T is made yet more negative, the familiar steps in resistance occurring at $h/2ne^2$ ($n = 1, 2, \ldots$) are observed due to the depopulation of 1D subbands in the bottom QPC. Finally, at $V_T = -3.8$ V, both channels are completely pinched off. Similar results are obtained when instead V_B is swept negative and $V_T = 0$.

These different regimes are evident in a waterfall plot of the conductance G of the device as a function of both V_T and V_B, shown in Fig. 3. Regimes (iii) and (iv) are now easily identified by their markedly different behavior. When only one wire is occupied (regime iv), uniform quantized steps in the conductance appear, indicating that transport is ballistic. This can be seen at the left and right sides of the base of the
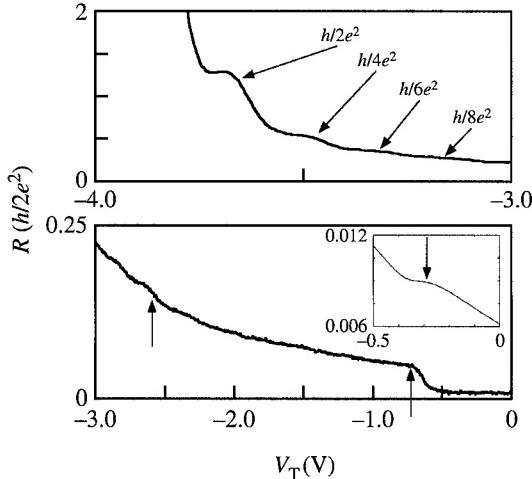
Fig. 2. Sample resistance versus V_T with $V_B = 0$. Inset at bottom shows transition from regime (i) to (ii) at $V_T = -0.3$ V. The transition from regimes (ii) to (iii) can be seen at $V_T = -0.7$ V while a distinct increase in slope at $V_T = -2.6$ V marks the transition from (iii) to (iv).
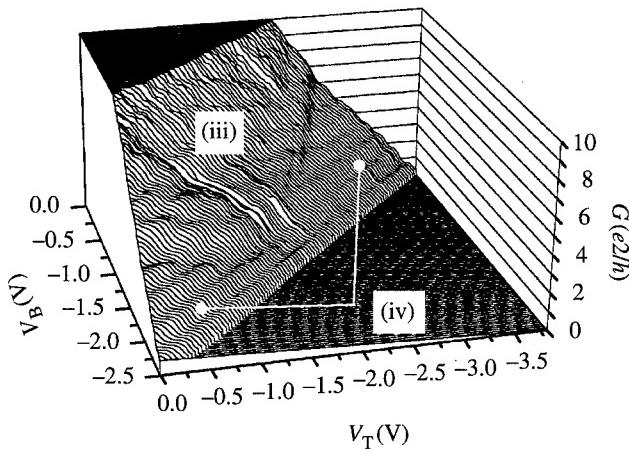
Fig. 3. Waterfall plot of conductance G versus V_T and V_B shows uniform quantized conductance steps for regime (iv), where only one 1D wire is present. A complex interference of the steps is evident in regime (iii), where both 1D wires are occupied.

plot. However, when both wires are occupied (regime iii), quantized steps are again present, but form a complicated interference pattern. By tracking the position of individual conductance steps as a function of both V_T and V_B, each step can be assigned to one quantum wire or the other. Clearly, each QPC width—and thus the number of occupied subbands—can be independently controlled by the action of V_T and V_B.

We now turn to the behavior of the conductance under an in-plane magnetic field B applied perpendicular to the direction of current flow. Figure 4 shows a contour plot of the logarithmic derivative of the resistance $\partial[\ln(R)]/\partial V_T$, as a function of V_T and V_B. The different regimes are again readily identifiable. Insets (a)–(e) show G versus B taken at several different values of V_T and V_B. Inset (a) was taken at $V_T = -0.25$ V and $V_B = 0.0$ V. The familiar 2D–2D anticrossing features, a conductance peak followed by a shallow dip, are centered about $B = 6.5$ T. This is consistent with the sample being in regime (i) [6].
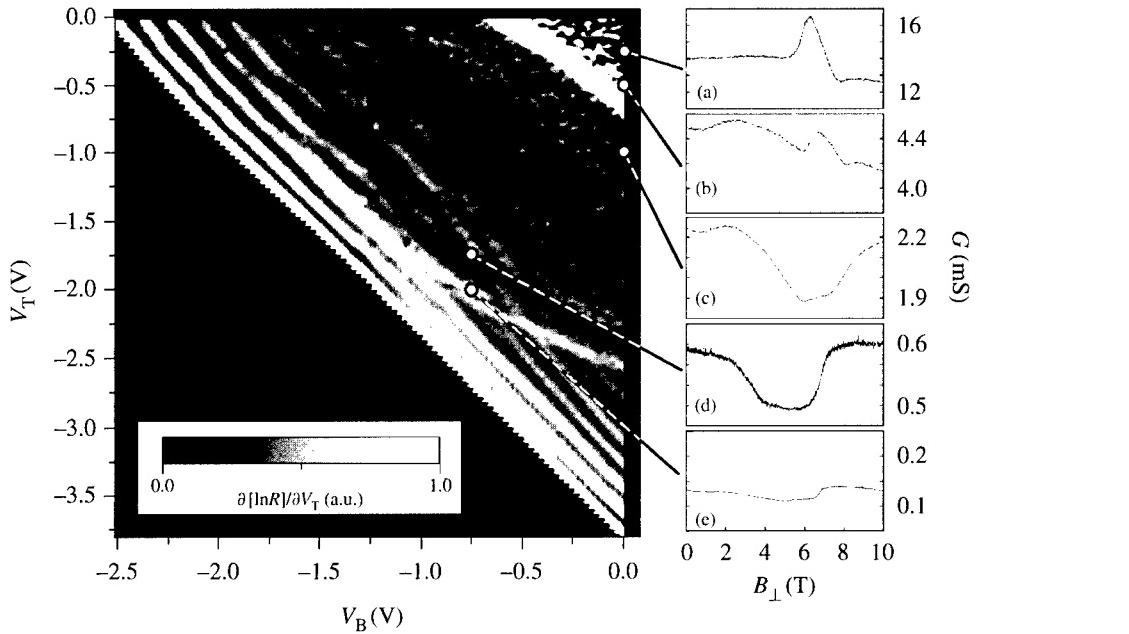
Fig. 4. Contour plot of the logarithmic derivative of the resistance $\partial[\ln(R)]/\partial V_T$. The different regimes are readily identifiable. Insets (a)–(e) are plots of G versus B taken at the indicated values of V_T and V_B. Here, where the in-plane B is perpendicular to the current direction, a broad conductance dip develops in regime (iii).

As the sample is biased into regime (ii), the 2D–2D conductance features begin to change. This is evident in inset (b), taken at $V_T = -0.50$ V and $V_B = 0.0$ V. Here the 2D–2D conductance features now appear to be superimposed upon a broad conductance dip. For $V_B = 0$ and values of V_T greater than -0.7 V, the 2D–2D conductance features *disappear completely* and are replaced by a *broad conductance minimum* slightly below 6.5 T. This feature is evident in insets (c) and (d), and is a hallmark of conductance behavior in regime (iii). Note that inset (d) is for $V_T = -1.75$ V and $V_B = -0.75$ V, showing that this conductance dip occurs over a wide range of V_T and V_B, so long as the sample is biased into regime (iii). Eventually, at high enough negative gate bias, the sample moves into regime (iv) where only one QPC remains occupied, and the conductance dip almost completely disappears. Inset (e), taken at $V_T = -2.00$ V and $V_B = -0.75$ V, is approaching this regime. This general behavior has been observed in several similar samples, indicating that this broad magnetoconductance dip is a characteristic feature of closely coupled ballistic 1D wires.

We note, however, that the broad conductance dip does *not* appear when the in-plane magnetic field B is oriented parallel to the current direction. This orientation was measured on a separate cool down of the same sample. Figure 5 shows the logarithmic derivative $\partial[\ln(R)]/\partial V_T$ versus V_T and V_B, similar to that of the earlier cooldown. However, application of a parallel B does not produce a broad conductance dip. Rather, small remnants of the 2D–2D anticrossing magnetoconductance features at $V_T = V_B = 0$ now persist into the 1D–1D regime, as shown in insets (a)–(e). These small peaks in conductance are attributed to the 2D–2D lead regions. (A change in the overall background magnetoconductance is unimportant for our purposes here.) Thus, the broad magnetoconductance dip is absent, and apparently is observed only for in-plane magnetic fields oriented perpendicular to the direction of current flow.

To discuss the origin of this 1D–1D magnetoconductance feature, it is useful to first consider previous work on the 2D–2D case [6]. The 2D–2D magnetoconductance features, exemplified in Fig. 4 inset (a), were shown to arise due to an in-plane magnetic field displacing the two 2D QW dispersion curves relative
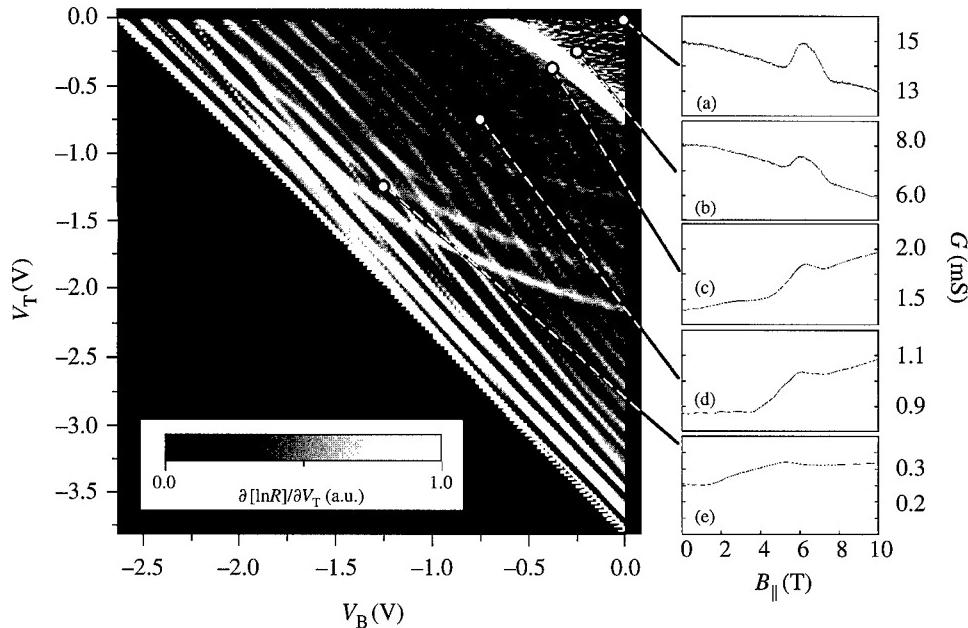
Fig. 5. Contour plot of the logarithmic derivative of the resistance $\partial[\ln(R)]/\partial V_T$, where again insets (a)–(e) are plots of G versus B taken at the indicated values of V_T and V_B. Here the in-plane B is oriented parallel to the current direction, and no broad conductance dip is seen to develop.

to each other in k-space by $\Delta k = eBd/\hbar$, where d is the distance between the two electron layers. At sufficiently high B the 2D dispersion curves anticross, forming a partial energy gap. Singularities in the 2D density of states and Fermi velocity occur at the upper and lower gap edges. As B is increased the gap edges cross the Fermi level, producing first a maximum in the conductance, followed by a minimum. These two magnetoconductance anticrossing features are centered about a point $B = [(2\pi n_1)^{1/2} + (2\pi n_2)^{1/2}]\hbar/ed$. Using the measured density values for our sample, we obtain $B \approx 6.5$ T, in agreement with the data of Fig. 4 inset (a).

A similar analysis based on the k-space offset of the 1D–1D dispersion curves, combined with the ballistic transport through the sample, explains the 1D–1D data. Due to tunnel coupling, at sufficiently high in-plane B perpendicular to the current, the 1D dispersion curves from one wire will anticross with their same-index counterparts from the other wire, opening a quasi gap of Δ_{SAS}, similar to the 2D case. The resulting dispersion curve thus consists of several anticrossed pairs of 1D subbands, as shown in Fig. 6. In our model, we assume that the transport is ballistic. Hence the conductance is given simply by the number of 1D subbands crossing the Fermi level. The application of B changes the conductance in two ways. (1) Whenever the Fermi level falls within a 1D anticrossing gap, the conductance is reduced by $2e^2/h$ from the uncoupled double quantum wire case. (2) Because for each pair of anticrossed 1D subbands there is a large increase in the density of states at the edges of the anticrossing gap, each anticrossed pair of subbands can accommodate many more electrons than if no anticrossing occurred. Since particle number is constant, as B is increased and the anticrossings are formed, the Fermi level *drops substantially* and intersects fewer 1D subbands, shown in Fig. 6B. This results in a large drop in conductance. Eventually B becomes large enough so that all the anticrossings in the 1D subbands (and the associated singularities in the DOS) rise above the Fermi level. The number of electrons accommodated by each subband thus correspondingly decreases. As a result, the Fermi level rises again (Fig. 6C) and many more subbands must then become occupied to accommodate all
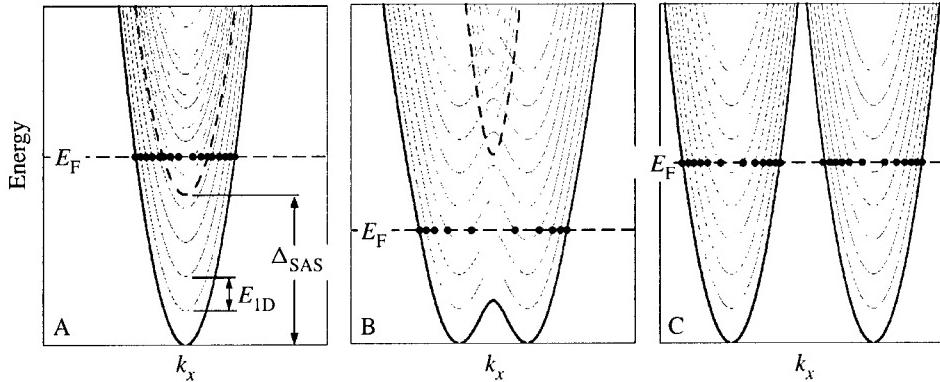
Fig. 6. Sketch of the dispersion curves for a symmetric pair of coupled double quantum wires, for the case where the subband spacing is much smaller than Δ_{SAS}. The in-plane magnetic field is applied perpendicular to the current direction. A, $B = 0$. B, $B \simeq 6.5$ T. C, $B =\sim 12$ T.

of the electrons, causing the conductance to rise again. While mechanism (1) will only be significant in QPCs with few occupied subbands, mechanism (2) will still be significant for relatively wide QPCs with dozens of occupied subbands. Only B perpendicular to the current will have this effect.

This behavior is consistent with that observed in the data of Fig. 4, and also with recent theoretical calculations by Lyo [7]. We note that a similar magnetoconductance dip has been observed by the Nottingham group in a mesa-etched double quantum wire, but was explained using an argument based on diffusive boundary scattering at the wire edges [2].

References

[1] N. Tsukada, A. D. Wieck, and K. Ploog, Appl. Phys. Lett. **56**, 2527 (1990).

[2] M. J. Gompertz, T. Ihn, P. C. Main, A. Nogaret, L. Eaves, M. Henini, and S. P. Beaumont, Physica **B249**, 162 (1998).

[3] K. J. Thomas, J. T. Nicholls, M. Y. Simmons, W. R. Tribe, A. G. Davies, and M. Pepper, Phys. Rev. **B59**, 12252 (1999); I. M. Castleton, A. G. Davies, A. R. Hamilton, J. E. F. Frost, M. Y. Simmons, D. A. Ritchie, and M. Pepper, Physica **B249**, 157 (1998).

[4] J. S. Moon, M. A. Blount, J. A. Simmons, J. R. Wendt, S. K. Lyo, and J. L. Reno, Phys. Rev. **B60**, 11530 (1999).

[5] M. V. Weckwerth, J. A. Simmons, N. E. Harff, M. E. Sherwin, M. A. Blount, W. E. Baca, and H. C. Chui, Superlatt. Microstruct. **20**, 561 (1996).

[6] J. A. Simmons, S. K. Lyo, N. E. Harff, and J. F. Klem, Phys. Rev. Lett. **73**, 2256 (1994).

[7] S. K. Lyo, Phys. Rev. **B60**, 7732 (1999).

Quantum confinement without walls

S. Tsujino[†], S. J. Allen, M. Thomas, T. Eckhause, E. Gwinn,
M. Rüfenacht

Center for Quantized Electronic Structures (QUEST), University of California, Santa Barbara, U.S.A.

J. P. Zhang, J. Speck

Materials Department, University of California, Santa Barbara, U.S.A.

H. Sakaki[†]

Institute of Industrial Science, University of Tokyo, Japan

We explore and demonstrate quantum confinement without walls in metal-clad InAs quantum wells. We observed intersubband absorption from confined states in InAs clad with Al, Sb, Nb, W, Pt, Ag, Au, Ti or In. We found that using this novel method, we can explore the physics and chemistry at the metal–semiconductor interface; reflection, autoionization, well-width fluctuation and interface reaction.

© 2000 Academic Press

Key words: metal–semiconductor interface, reflection, autoionization, intersubband absorption.

The metal–semiconductor interface is an important element of nearly all semiconductor electronic and photonic devices forming Schottky or ohmic contacts. Although they have been studied for many decades, there are few experiments that explore the quantum mechanical boundary conditions. Nearly every aspects of the band structure is discontinuous at the interface and quantum mechanical reflection and transmission are difficult to anticipate *a priori*. The boundary condition will become more important in nanoscale semiconductor devices, since the ratio of contacting area to the volume of the active region of the device becomes larger as the size of the device becomes smaller. The boundary condition is also important for quasi-particle transport in superconductor–semiconductor devices [1] and for spin transport in ferromagnet–semiconductor devices [2, 3].

In this work, we demonstrate and explore quantum confinement of carriers in InAs quantum wells defined by a metal–quantum well interface (Fig. 1). Since the InAs–metal interface forms ohmic contact, it is usually assumed that the interface is transparent and quantum mechanical reflection can be ignored. It seems clear that this is not likely and the quantum mechanical transport through the interface will be sensitive to quantum mechanical reflection and transmission, and material reactions at the interface. While the metal presents no barriers or walls, band structure discontinuities can lead to quasi-bound states and *confinement without walls.*

[†]Also at: Quantum Transition Project, Japan Science and Technology Corp., Japan.
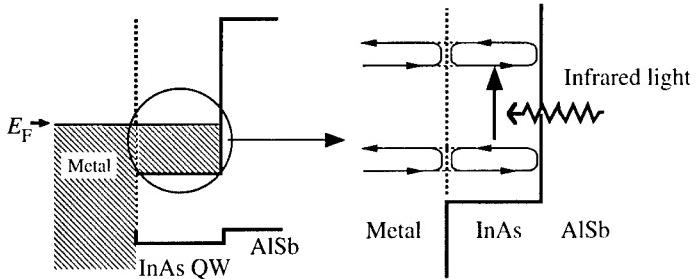
**Fig. 1.** Confinement without walls in metal-clad InAs structures.

The *intersubband* optical absorption spectrum should be sensitive to the reflection and the quantum mechanical boundary condition at the metal–semiconductor interface, since the magnitude and the phase of the reflection coefficient determines the width and the frequency of the resonance. Therefore, we explore intersubband transitions in InAs quantum wells terminated by various metal–InAs interfaces and produced in various ways. While this approach has the potential to test physical models of the metal–semiconductor interface, systems that support well-defined quantum states are potentially important for ultra-fast resonant tunneling diodes and infrared detectors.

The metal-clad quantum wells are based on InAs–AlSb structures and were prepared by *in-situ* MBE growth and by *ex-situ* processing. *Ex-situ* samples are perhaps most interesting and prepared as follows. We start from the InAs–AlSb QW structure, remove the top AlSb barrier and GaSb cladding layer by selective wet chemical etching in ambient environment. Immediately after the etching process, we put the sample into a metal deposition chamber, evacuate down to $10^{-6} \sim 10^{-7}$ torr, and deposit a metal film by sputtering or electron beam evaporation. For some samples, we cleaned the InAs surface by Ar plasma by removing a few nanometers of the material including oxide. We used modulation doped 15 nm thick InAs–AlSb QWs. The electron concentrations are nominally 7.5×10^{12} cm^{-2} and the mobilities are 6 to 15 m^2 (Vs)$^{-1}$. Using this method, we investigate the confinement of electrons in InAs terminated by Al, W, Pt, Ag, Nb, In, Ti, Au, or Ni. *In-situ* samples were prepared by growing the structures by molecular beam epitaxy method including the metal layer (Al, Sb, or Nb).

We measured the mid-infrared absorption spectrum of the samples using Fourier transform infrared spectrometer (FTIR) in multi-pass waveguide geometry with 5 mm length and 500 μm thickness. The absorption is measured with TM polarized light at 300 K. The optical reference is provided by the sample prepared similarly but with the InAs layer removed.

In-situ samples should have a sharper and cleaner interface and are a good benchmark for metal-clad quantum wells. *Ex-situ* samples suffer more drastic processing [4], but recovery of well-defined intersubband absorption opens a richer arena in which to correlate quantum states and interface physics and chemistry. We will focus on the *ex-situ* experiments in the following.

Figure 2 shows the absorption spectrum of Al-clad InAs structures. The curve A is the intersubband absorption of the original InAs quantum well clad by AlSb barriers. The sharp absorption from the first excited subband to the second excited subband is observed at 176 meV. The material was then processed to create an Al-clad structure. After chemically exposing the InAs well, Al was deposited by sputtering after RF plasma cleaning for 30 s with plasma potential about 100 V. We found that a clean absorption was still observed (curve B of Fig. 2), showing that the Al–InAs interface is very reflective for electrons [5]. We identify the observed absorption of the Al–InAs structures as the transition from the first excited to the second excited subband, but note the peak shift and broadening. The peak shift is 29.8 meV toward higher photon energy and the increase of width Δ_{Al} is 30.5 meV. This observation is explained by thinning of InAs of about 2 nm
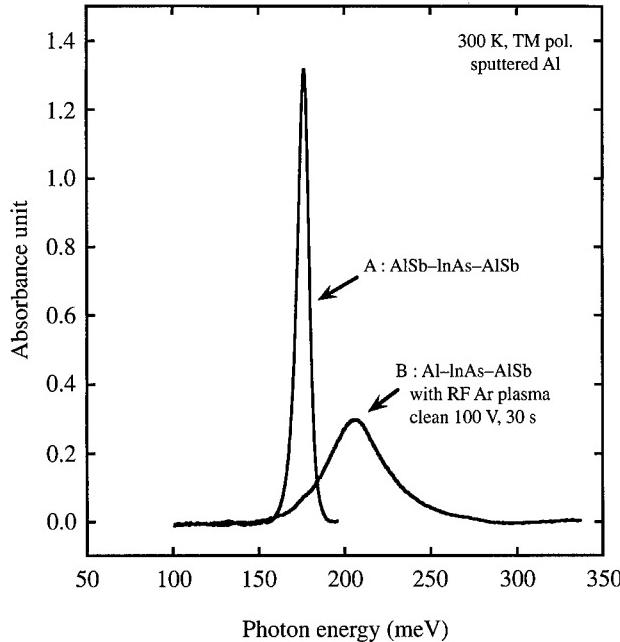
Fig. 2. Absorption spectrum of the Al-clad InAs structure. A and B show the intersubband absorption of the same sample before (A) InAs quantum states are confined by AlSb barriers and after (B), processing the sample to the Al–InAs structure.

and a well-width fluctuation induced by the process and/or by reaction of Al with InAs [6]. The data suggest that despite the expected transmissivity of the Al–InAs interface, the electrons are strongly reflected either by band structure discontinuity or formation of a AlAs barrier.

In Fig. 3, we show the absorption of a W–InAs quantum well compared with the barrier bound system. The process followed the Al–InAs quantum well shown in Fig. 2. We found that W–InAs exhibits intersubband absorption, shifted toward higher energy and broadened, showing that the W–InAs interface is also reflective in a way somewhat similarly to Al–InAs. The amount of peak shift, 26.1 meV, is of the same order as the shift in Al–InAs. However, W–InAs shows an anomalous absorption in the lower energy side of the original intersubband peak which is not observed in the Al–InAs structure. The width increase $\Delta_W = 67.1$ meV is more than factor of 2 of Δ_{Al}. This anomalous absorption cannot be accounted for by well-width fluctuation. Similar anomalous low-frequency absorption is also observed in structures using Pt, Ag, In, Au, Nb, and Ti.

Nearly same peak shift of both Al–InAs and W–InAs indicates that the effect of the etching and Ar plasma cleaning to the InAs film are similar for both samples. Also cross-sectional transmission electron microscope study of the W–InAs structure reveals that the W–InAs interface is abrupt (the interface transient is less than 2 ML, which is equal to the resolution limit of the instrument). These observation indicates that in the W–InAs structure, the electron states in the InAs may autoionize as Fano described some time ago [7]. In our case, the InAs states are autoionized by the coupling to the metallic electronic states as a continuum.

Based on these result, we estimated the reflection coefficient R at the W–InAs interface. Using the processed Al–InAs quantum well absorption as a base line, we estimate the increased absorption due to autoionization. We calculate the broadening of the autoionized subband as a function of R. The line width increase indicates that $|R|^2 \simeq 0.47$. The phase of R at W–InAs is estimated to be close to π, since the peak shift of W–InAs is similar to that of the Al–InAs structure. We found a similar broadening in the Nb–InAs interface [8] to W–InAs, which indicates that quasi-particle transport across the interface is affected by the interface transmission. This is consistent to the observation that single Andreev reflection probability is rather low ($\sim$3%) [9].
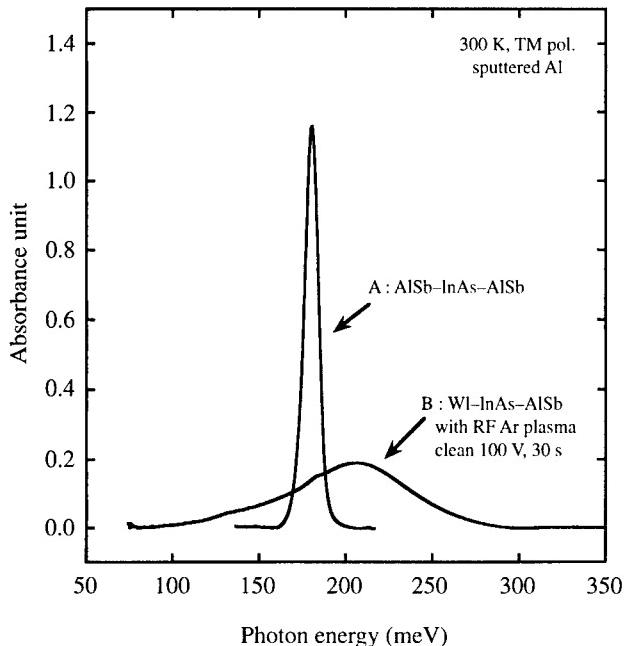
Fig. 3. Absorption spectrum of the W-clad InAs structure. A and B show the intersubband absorption of the same sample before (A) InAs quantum states are confined by AlSb barriers and after (B), processing the sample to the W–InAs structure.

In conclusion, we successfully observed intersubband absorption in metal-clad InAs structures using various metals. We showed that intersubband spectroscopy can sensitively detect the metal-semi interface conditions and is promising to explore the rich physics and chemistry at the interface; reflection, autoionization, well-width fluctuation, and interface reaction.

References

[1] A. Kastalsky, A. W. Kleinsasser, L. H. Greene, R. Bhat, F. P. Milliken, and J. P. Harbison, Phys. Rev. Lett. **67**, 3026 (1991).

[2] D. D. Awschalom, Surfaces and Interfaces of Mesoscopic Devices, SIMD '99, Maui, U.S.A., 1999.

[3] G. Schmidt, R. Fiederling, G. Reuscher, W. Ossau, A. Waag, and L. W. Molenkamp, Superlatt. Microstruct. **27**, 297.

[4] P. H. Magnee, S. G. den Hartog, B. J. van Wees, T. M. Klapwijk, W. van de Graaf, and G. Borghs, Appl. Phys. Lett. **67**, 3569 (1995).

[5] M. V. Weckwerth, J. P. A. van der Wagt, and J. S. Harris, Jr., J. Vac. Sci. Technol. **B12**, 1303 (1994).

[6] M. Missous, E. H. Rhoderick, and K. E. Singer, J. Appl. Phys. **59**, 3189 (1986).

[7] U. Fano, Phys. Rev. **124**, 1866 (1961).

[8] T. A. Eckhause, S. Tsujino, K. W. Lehnert, E. G. Gwinn, S. J. Allen, M. Thomas, and H. Kroemer, Appl. Phys. Lett. **76**, 212.

[9] C. Nguyen, H. Kroemer, and E. L. Hu, Appl. Phys. Lett. **65**, 103 (1994).

Directed self-assembly process for nano-electronic devices and interconnect

EDWIN C. KAN, ZENGTAO LIU

School of Electrical and Computer Engineering, 404 Phillips Hall, Cornell University, Ithaca, NY 14853, U.S.A.

Requirements for nano-electronic devices and interconnect in conventional logic circuit architecture are examined to reveal the possible device structure containing self-assembly (SA) features. For integration with lithography, we discuss the features of the SA process and then present a full process flow for a novel type of SA nano-crystal EEPROM cells. The logic functionality below the lithography limit derived from the SA dot arrays and the I–V characteristics including the Coulomb blockade effect are then presented.

© 2000 Academic Press

Key words: self-assembly, nanoelectrics, CMOS technology.

1. Introduction

Electronic functions such as multi-bit nonvolatile memory cell, fault-tolerant quad logic supplement, flexible FPGA (field programmable gate array) router, electronic synapse, and tera-bit mass storage can be realized by combination of the self-assembly (SA) process of the floating nano-crystals and the conventional CMOS lithography steps [4]. The regular or irregular SA nano-crystal array formulated by minimization of surface strain energy on solid interfaces [5] can have resolution below the lithography limit, and can hence take advantage of the whole-wafer nature of formation and the Coulomb blockade effects [8]. The basic device structure proposed here is similar to the conventional EEPROM with tunneling oxide and a top writing gate (see Fig. 1). The floating gate is replaced by the nano-crystal array, and the source/drain contacts can lead signals out in all directions. The new geometrical arrangement significantly enhances the functional density and provides new design flexibility in EEPROM scaling. Figure 2 illustrates the influence of physical forces on the size and distance distribution of the SA nano-dots. However, array of SA nano-crystals can provide useful electronic functions only if combined with arbitrary design patterning, such as UV lithography [2], nano-imprinting [3] and E-beam [9]. These arbitrary design patterns, which are also affected by similar surface forces, are used to formulate interconnect for signals processed or stored in the nano-dot arrays. Perturbation of the SA process caused by surface roughness, pre-registered patterns and long-range forces has been studied [7]. Our proposed structure has an emphasis on CMOS technology compatibility. Sample SEM pictures of the SA processes used for characterization are shown in Figs 3 and 4, where Au, Cr, Si, SiO_2 and Si_3N_4 are employed. The nano-crystal formation is also heavily dependent on the thickness of the starting wetting layers, especially when the Ostwald ripening (with atomic transfer by surface diffusion instead of by gas phase) dominates or when the surface mobility is low.

A SEM picture of self-assembled irregular arrays of nano-dots showing four operational quadrants.

Self-assembled regular or irregular array of nano-dots. An irregular cluster of dots can be thought of as one operational dot.
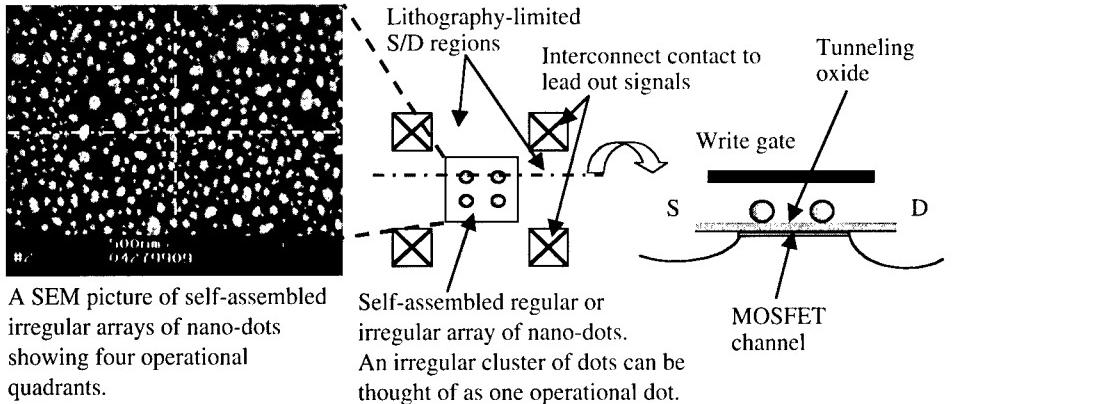
Fig. 1. A basic quad-element derived from an EEPROM device with self-assembled nano-crystal floating gate. Notice that the nano-crystals do not need to be evenly distributed for the four-quadrant applications. They only need to be small enough to take advantage of the Coulomb blockade effect.
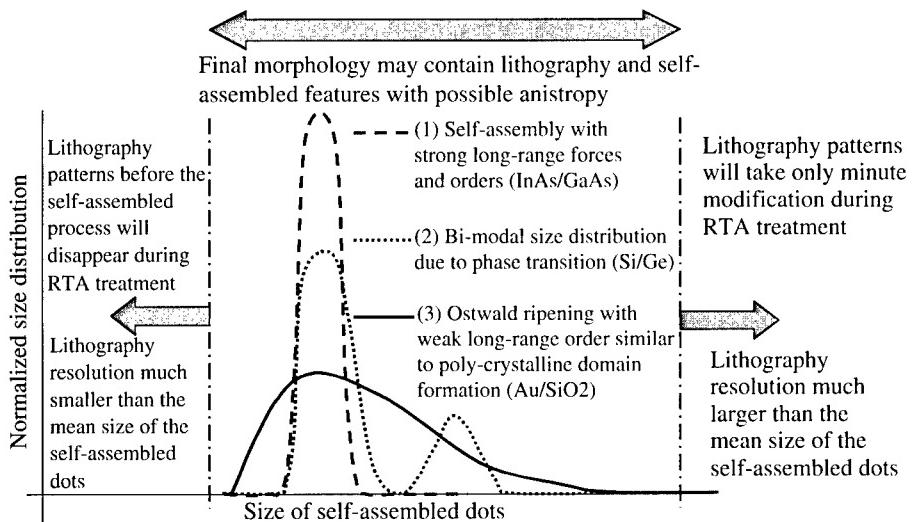
Fig. 2. Size distribution of the self-assembled dots and its relations with the lithography process. Notice that the layer thickness dependence has not been taken into consideration.

2. Functional density of the SA structure

In present CMOS technology, if the active devices are mostly contained in one layer, circuit modules based on conventional architecture generally obeys the extended Rent's Rule [6] on interconnect density. A generalized version can be expressed as $P = K \times G^{\beta}$, where P is the number of interconnect on the perimeter of the circuit module, K is the average number of I/O per logic gate, G is the number of logic gates, and β varies between 0.1 and 0.7. If the SA structures perform similar logic functions in the same active layer of the CMOS environment, the pitch of the perimeter interconnect will be equal to the SA features when $\beta = 0.5$ and $K = 4$, and can be larger than, but still comparable to, the SA feature even for small β. The Rent's rule renders a lower limit on interconnect pitch leading out of a conventional circuit architecture. For electronic functions based on SA structures, if the functional density is much smaller than lithography-limited CMOS, the interconnect pitch and CMOS interface drive capability will both be problematic. If the
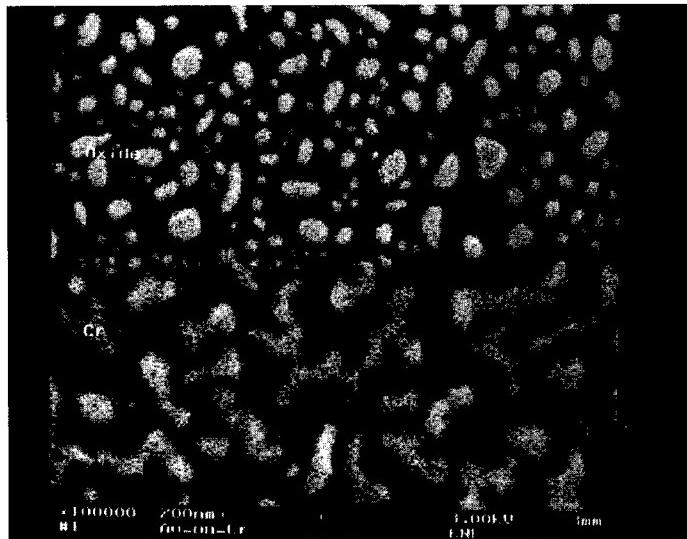
Fig. 3. Self-assembled Au dots on adjacent oxide and Cr regions. The direct comparison can isolate the thermal effects and offer information on formation energy and effects of surface roughness.
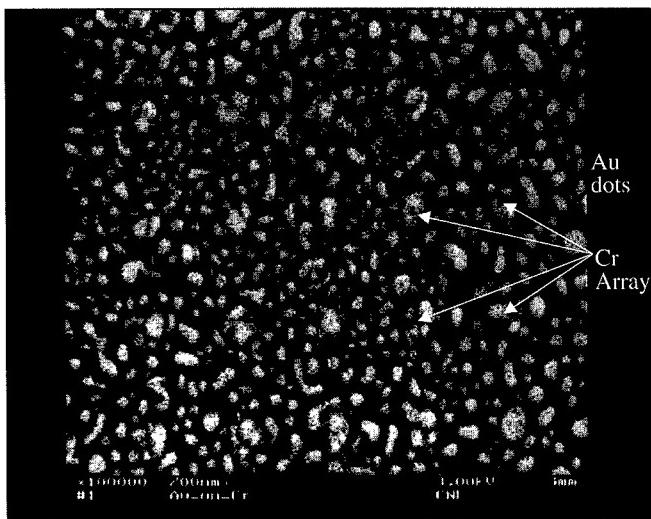
Fig. 4. Self-assembled Au dots on smooth oxide with E-beam-prescribed Cr arrays. This hybrid process can be used to characterize self-assembly forces since Cr has much smaller surface mobility than Au and serves as place-holders.

functional density is only comparable to CMOS, then the SA process can only enhance structural stability. This leaves a small window of functional density of the SA structures if interface with CMOS circuitry is necessary. The number of zones of nano-crystals floating gates associated with electronic functions in the proposed nano-crystal EEPROM cell is hence quite limited, unless direct information access can be provided from the third dimension as in the tera-bit storage case [4]. Another possible way to surpass the Rent's limit is by considering that the electronic functions are very different in nature than logic circuits, as in the case of electronic synapse [1]. This, however, implies a complete change in processing architecture. If only four
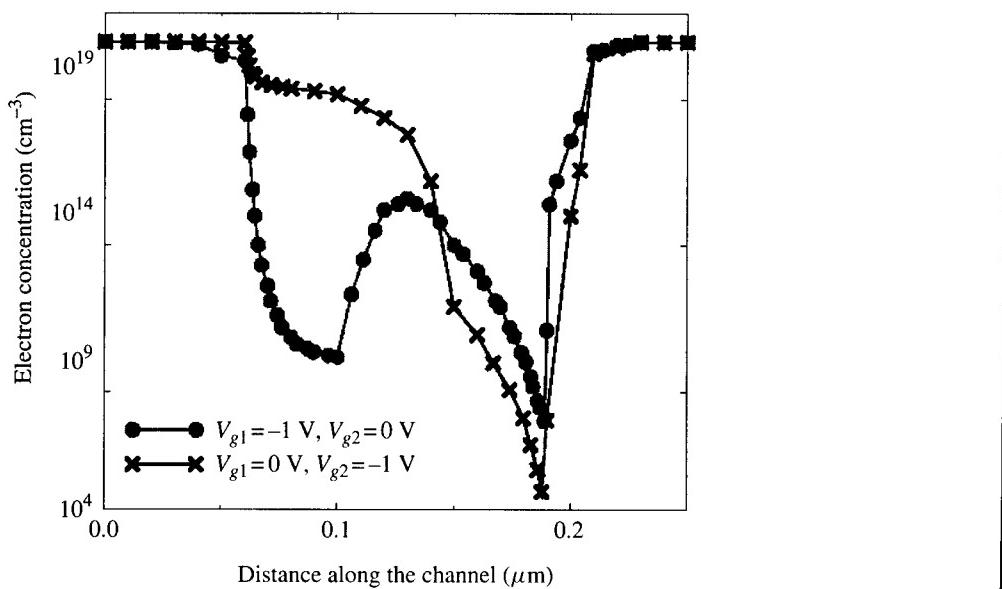
Fig. 5. Simulated channel electron concentration under a split-gate MOSFET. The two gates can be easily distinguished by DIBL asymmetry.
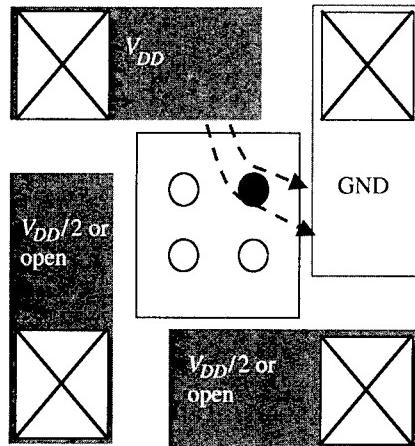

Fig. 6. To write an individual dot using the peripheral contacts, two corner contacts are fully biased, while the other two half biased or floating.

zones are used in the nano-crystal floating gate as in our four-bit EEPROM cell, the logic function is led out of the SA structure from the field effect in the floating charges.

3. Device operations of the novel EEPROM cell

In the above applications, in order to properly access each zone of nano-crystals from the peripheral contacts, it is necessary that the drain-induced barrier lowering (DIBL) effect in the MOSFET structure is strong so that the position of each zone along the channel can be distinguished by terminal I–V characteristics. Fig-
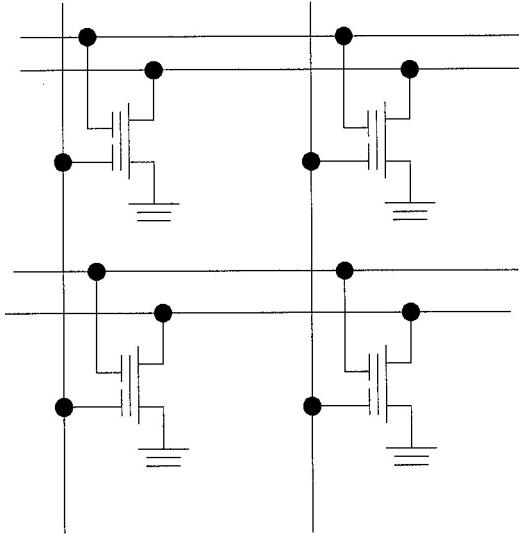
Fig. 7. Equivalent circuit of the electron synapse with output set by the convolution of the channel current history stored in nano-crystal floating gates.

Step 1: field oxide and active region (Mask 1)

Step 2: tunneling oxide growth and nano-dot floating gate formation

Step 3: PECVD control oxide and sputtered poly/silicide formation

Step 4: control gate patterning (Mask 2)

Step 5: spacer formation (dielectric deposition and etch)

Step 6: ion implantation for source and drain regions

Step 7: contact hole formation (Mask 3, control gate contact not shown)

Step 8: interconnect patterning (Mask 4, control gate interconnect not shown)

	Si		oxide
	metal		poly

Fig. 8. Demonstration of process flow for the nano-crystal floating gate device. Notice that this flow is for illustration purposes only. The geometry is not drawn to scale.

ure 5 shows the simulated channel carrier concentration in a 50 nm split-gate MOSFET using a conventional drift-diffusion model, where DIBL asymmetry is indeed significant. Programming of each individual zone can be achieved through hot-carrier injection with the bias condition shown in Fig. 6. Erasing all four nodes can be achieved by tunneling through the control gate. This is similar to the popular FLASH operations [6].
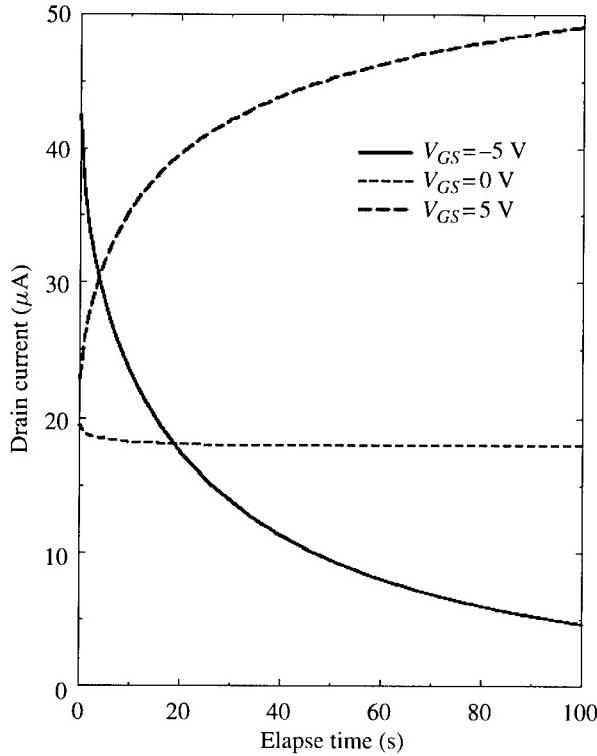
Fig. 9. Preliminary measurement of the Au nano-crystal EEPROM cell. The drain bias is 5 V. The solid line has control-gate bias at −5 V, the short-dashed line has 0 V and the long-dashed line has 5 V. The drain current is tuned by the floating charge in the nano-crystals injected by hot channel carriers. Leakage is almost negligible with no control gate bias. Hot carrier charging is self-limited.

Since the size of the nano-crystals is much smaller than the zone dimension, the Coulomb blockade effect on the charge quantization of each nano-crystal will still be strong, and the threshold hysteresis can be observed as [8]: $\Delta V_T = \frac{qvn_{dot}}{\varepsilon_{ox}\varepsilon_0}\left(t_{cntl} + \frac{1}{2}\frac{\varepsilon_{ox}}{\varepsilon_{dot}}t_{dot}\right)$, where v is the number of elemental charges on one nano-crystal, n_{dot} is the average number of nano-crystals in each zone, t_{cntl} is the thickness of the control oxide, and t_{dot} is the thickness of the nano-crystal. Figure 7 shows the equivalent circuits for the synapse operations of the four-quadrant EEPROM cells.

The process flow illustrating the necessary steps for fabrication of the proposed EEPROM cell is shown in Fig. 8. Special attention has been paid to CMOS process compatibility. It can be observed that the deviation from conventional CMOS processes is small, mainly in the formation of control gate by PECVD (plasma-enhanced chemical vapor deposition) oxide due to a more stringent thermal budget. Full electronic characteristics of the novel devices are not provided here due to length limitation. Figure 9 shows the hot-carrier charging and data retention characteristics of the fabricated cell with Au nano-crystals under different gate biases. Notice that even though the process is not yet optimized for EEPROM operations as nonvolatile memory, the large difference in charging and retention time strongly indicates that the new devices are feasible for the functional circuits described above.

References

[1] C. Diorio, P. Hasler, B. A. Minch, and C. A. Mead, IEEE Trans. Elec. Dev. **43**, 1972 (1996).

[2] T. I. Kamins and R. S. Williams, Appl. Phys. Lett. **71**, 1201 (1997).

[3] T. I. Kamins, D. A. A. Ohlberg, R. S. Williams, W. Zhang, and S. Y. Chou, Appl. Phys. Lett. **74**, 1773 (1999).

[4] E. C. Kan and Z. Liu, Self-assembled nano-particles as electronic devices, in *Proceedings of the Sixth Foresight Conf. on Molecular Nanotechnology* (Santa Clara, CA, 1997).

[5] R. Notzel, N. N. Ledentsov, L. Daweritz, M. Hohenstein, and K. Ploog, Phys. Rev. Lett. **67**, 3812 (1991).

[6] J. M. Rabaey, *Digital Integrated Circuits—a Design Perspective* (Prentice Hall, New Jersey, 1996).

[7] Z. Suo and Z. Zhang, Epitaxial films stabilized by long-range forces, Phys. Rev. **B58**, 5116 (1998).

[8] S. Tiwari, F. Rana, K. Chan, L. Shi, and H. Hanafi, Single charge and confinement effects in nanocrystal memories, Appl. Phys. Lett. **69**, 1232 (1996).

[9] J. Y. Yew, L. J. Chen, and K. Nakamura, Epitaxial growth of $NiSi_2$ on (111)Si inside 0.1–0.6 μm oxide openings prepared by electron beam lithography, Appl. Phys. Lett. **69**, 999 (1996).

Electron-beam exposure of self-assembled monolayers of 10-undecenoic acid

M. N. KOZICKI, S.-J. YANG, B. W. AXELROD

Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85287-5706, U.S.A.

Tapping mode atomic force microscopy and capacitance versus voltage measurements were employed to study the effects of electron-beam exposure on self-assembled monolayers of 10-undecenoic acid. It was established that exposure increases chemical/mechanical stability, resulting in a thicker layer following a solvent treatment designed to remove residual monomers. Electron exposure also reduces the effects of pinholes in the monolayer, thereby improving dielectric quality.

© 2000 Academic Press

Key words: undecenoic acid, monolayer, electron-beam.

Self-assembled monolayers (SAMs) such as thiols on gold, siloxanes on SiO_2, and carboxylic acids on SiO_2 or Al_2O_3 have been studied extensively for possible applications in electronic device fabrication. Successful examples of the use of SAMs as lithographic resists [1] and as dielectric materials [2–4] have demonstrated their potential in this context. Lithographic applications of SAMs have already been explored by our group and the results reported elsewhere [5]. In our present work, 10-undecenoic acid, (UA), $H_2C = CH(CH_2)_8COOH$, self-assembled on SiO_2 was examined to assess changes in its chemical/mechanical and electrical properties following electron-beam exposure. The motivation for this study was to assess the suitability of electron-beam exposed UA as a self-assembling patternable nanoscale dielectric layer. Our previous work showed that electron-beam exposure of UA results in a crosslinked matrix, the linkages being formed between adjacent molecules at the carbon double bond in the backbone [5]. Certain solvents may be used to dissolve the unexposed monomers but the exposed crosslinked SAM has significantly reduced solubility and therefore remains on the surface following solvent treatment; UA is therefore patternable by e-beam exposure. For this study, tapping mode atomic force microscopy (TMAFM), a noncontact method which is ideal for probing soft materials, was employed to examine the mechanical properties (specifically the apparent thickness/rigidity) of the exposed pattern. Metal/dielectric/metal structures were used in capacitance versus bias voltage analysis to assess dielectric properties.

For the TMAFM study, oxidized silicon wafers were used to support the SAM. The sample cleaning, oxidation, and SAM deposition techniques have been described elsewhere [5]. A converted scanning electron microscope equipped with ELPHY® lithographic controls was used for exposure. A methanol dip was employed to remove the UA monomers in the unexposed regions. A Digital Instruments Nanoscope IIIa TMAFM was then used to examine the residual patterns. The results of this experiment are shown in Fig. 1. Thickness measurement of this type of layer by TMAFM poses some uncertainty due to changes in the
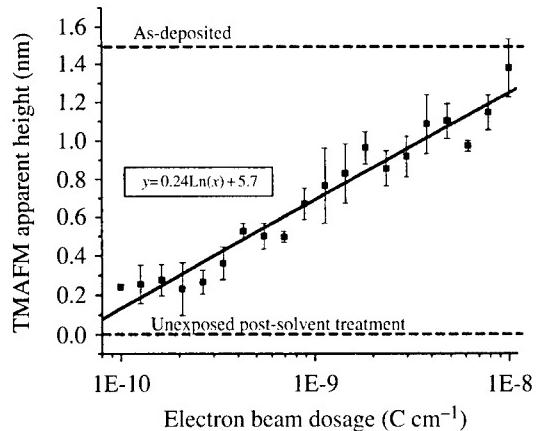
Fig. 1. TMAFM analysis of exposed 10-undecenoic acid features as a function of dose.

nature of the surface during exposure [6], however, the relative height from one pattern to another may be determined with a high degree of confidence. Figure 1 clearly shows the general decreasing trend of apparent height from the as-deposited (no solvent treatment) value as the dose decreases. The logarithmic relation is analogous to the response of negative resist following UV exposure and development. This is not surprising as larger doses will result in a higher degree of crosslinking which will, in turn, decrease solubility in the methanol and also increase the mechanical 'rigidity' of the layer. Both effects will lead to a thicker pattern as determined by TMAFM following exposure and solvent treatment.

Metal/SAM/SiO_2/metal and metal/SiO_2/metal (control) samples were fabricated in order to perform the capacitance versus voltage measurements. First, 300 nm of aluminum was evaporated onto oxidized silicon substrates to form the (blanket) bottom contact. Next, 12 nm of SiO_2 was deposited by remote plasma chemical vapor deposition (RPCVD) at 350 °C using N_2O + SiH_4 reactants. Half of each sample was dipped into 20 : 1 buffered HF for 20 s to remove the RPCVD SiO_2 and expose the bottom contact. After this, the samples were ultrasonically cleaned in acetone for 10 s. For the SAM-coated samples, a standard monolayer deposition process was used [5]. Some samples were then exposed with the electron beam with a dose of approximately 10 μC cm^{-2} to create 1.5 nm thick films. A shadow mask was used to deposit aluminum onto the sample surface for the top contact. The top contact was used for voltage biasing, the bottom contact was set to ground, and measurements were performed in a light-tight box. In order to explain the dielectric properties of the metal/SAM/SiO_2/metal structures, we used an equivalent model consisting of two parallel paths, as shown in the insert in Fig. 2A. The first path contains the capacitance of an ideal SAM in series with the capacitance of the SiO_2 layer, which was measured independently (using the control structure) and is frequency independent. The second path is a leakage pathway based upon the pinhole leakage model discussed by Jin *et al.* [4] and Vuillaume [3]. The pinholes are characterized by $R_{Leakage}$ and $C_{Leakage}$ which determine a time constant for the electron to travel through defects in the SAM. The effect of leakage through the pinholes is strongly frequency dependent. When a constant voltage is applied no current should flow; we have taken I–V measurements that show this to be the case. At high frequencies the leakage pathway becomes negligible because the polarity switches before an electron has time to traverse the SAM. The onset frequency of the effects from the leakage pathway is a good indicator of the quality of the SAM. The capacitance versus voltage dependence on frequency was determined by measuring the capacitance versus voltage characteristics of the Al/SAM/SiO_2/Al structure over one period at a range of frequencies from 1 kHz to 1 MHz. Figure 2A shows the average capacitance over one period versus frequency. At high frequencies the capacitance converges to a constant value determined by the series capacitance of the SAM and the SiO_2,
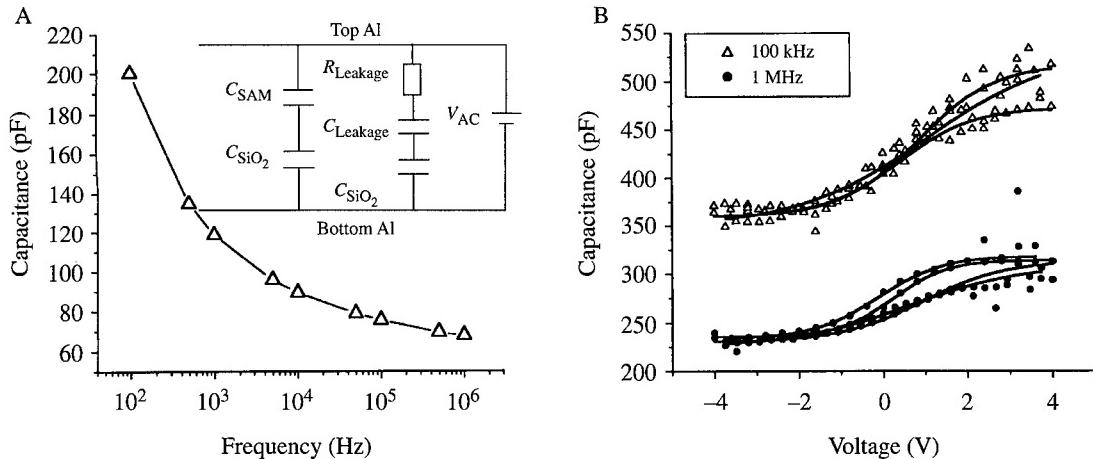
Fig. 2. A, Capacitance versus frequency for metal-oxide-SAM-metal structures without electron-beam lithography exposure. The capacitance is taken from the average capacitance over one period. B, Capacitance versus voltage over one period at 100 kHz and 1 MHz for an exposed metal-oxide–SAM–metal structure. Four measurements at each frequency are shown each fitted to a Boltzmann distribution.
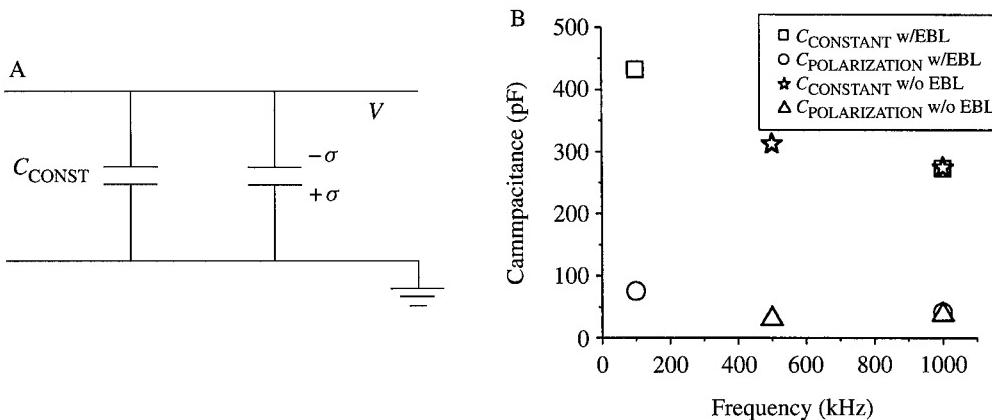
Fig. 3. A, Equivalent model circuit for the SAM. C_{CONST} is the constant capacitance and $+\sigma/-\sigma$ denote the location of the effective bound charge due to the intrinsic polarization of the SAM. B, Values of the constant capacitance of the SAM and voltage dependent capacitance, due to the intrinsic polarization of the SAM, versus frequency for SAMs with and without electron-beam lithography processing.

thus the capacitance of the SAM can be measured directly. We assume that the capacitance of the SiO_2 layer is not affected by the addition of the SAM. Figure 2B shows the capacitance of the SAM versus voltage over a single period at 100 kHz and 1 MHz, the data are well fitted by a Boltzmann distribution.

Similar polarity-dependent capacitance has been observed for metal/SAM/semi-conductor structures made with orthophenanthroline [2]. The success of the Boltzmann fit indicates that the leakage pathways are having a negligible effect. For SAMs that have been exposed to electron beam the Boltzmann distribution is present at frequencies as low as 100 kHz, whereas, the SAMs that have not been exposed to electron beam do not show the Boltzmann distribution until 500 kHz. This indicates that the pinholes in the electron beam exposed samples are less conductive and fewer in number than in the samples that have not been exposed to electron beams. The voltage dependence of the SAM capacitance is due to an intrinsic polarization of the SAM which can be modeled using a bound charge construct [7]. The capacitance versus voltage data show

that there is also a constant component to the capacitance of the SAM. The constant capacitance and the voltage-dependent capacitance can be modeled by a constant capacitance in parallel to two plates containing the bound charge created by the polarization of the SAM, as shown in Fig. 3A. When the SAM is positively biased the negative bound charge will offset some of the charge deposited on the capacitor plates, thereby increasing the capacitance until the bound charge becomes negligible in comparison with the total charge on the capacitor; the reverse will occur when negatively biased. The Boltzmann fit to the data can be used to determine values for the constant capacitance and the magnitude of the voltage-dependent capacitance. Figure 3B shows values for the constant capacitance and the voltage-dependent capacitance at a range of frequencies for both electron beam exposed SAMs and SAMs without electron-beam exposure. The electron-beam exposure does not appear to affect the constant capacitance of the SAM nor the voltage-dependent capacitance (supported by AM1 molecular orbital calculations which indicate that exposure should change the molecular dipole moment by less than 2.5%). It is not possible to determine from this data if the constant capacitance is frequency dependent or if the apparent frequency dependence is a result of the onset of effects from the leakage pathway.

References

[1] L. Stockman, G. Neuttiens, C. Van Haesendonck, and Y. Bruynseraede, Appl. Phys. Lett. **62**, 2935 (1993).
[2] J. Tanguy, Thin Solid Films **13**, 33 (1972).
[3] D. Vuillaume, Organic insulating films at nanometer scale, in *Proceedings of the Amorphous and Crystalline Insulating Thin Film*, edited by W. L. Warren, R. A. B. Devine, M. Matsumura, S. Cristolveanu, Y. Homma, and J. Kanicki (Material Research Society, Boston, MA, Vol. 446, 1996) pp. 79–90.
[4] Z. H. Jin, D. V. Veznov, Y. W. Lee, J. E. Zull, C. N. Sukenik, and R. F. Savinell, Langmuir **10**, 2662 (1994).
[5] T. K. Whidden, S.-J. Yang, A. Jenkins-Gray, M. Pan, and M. N. Kozicki, J. Electrochem. Soc. **144**, 1337 (1997).
[6] S. J. T. Van Noort, K. O. Van der Werf, B. G. De Grooth, N. F. Van Hulst, and J. Greve, Ultramicroscopy **69**, 117 (1997).
[7] D. J. Griffiths, *Introduction to Electrodynamics* (Prentice Hall, New Jersey, 1989) pp. 158–195.

Nanoscale effects in devices based on chalcogenide solid solutions

M. N. Kozicki, M. Yun, S.-J. Yang, J. P. Aberouette, J. P. Bird

Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85287-5706, U.S.A.

Solid solutions of metals such as silver in arsenic trisulfide exhibit a variety of interesting characteristics, including the ability to bring the metal out of solution by electrodeposition. This allows voltage-controlled switching characteristics to be realized in two terminal devices which may be fabricated to nanoscale dimensions. In addition, surface electrodeposits may be formed and subsequently broken at grain boundaries to create structures that have adjustable current–voltage characteristics at low temperatures. This paper highlights some initial results of the characterization of nanoscale structures based on such solid solutions.

© 2000 Academic Press

Key words: chalcogenide, solid solution, electrodeposition.

Metals such as silver or copper may be dissolved in chalcogenide glasses such as arsenic sulfide or germanium selenide to form solid solutions [1, 2]. These solid solutions have a relatively high resistivity (as compared with the solid metal) as room temperature conduction is dominated by ion transport through the solid electrolyte. If electrodes are formed in contact with a layer of the solid solution and a voltage is applied between them, the positively charged metal ions migrate toward the cathode region. If the anode contains the same metal as that in solution, the metal will dissolve into solution at the anode as it comes out of solution at the cathode. This electrically stimulated reduction–oxidation reaction results in the formation of a stable metallic electrodeposit which may be made to extend completely from the cathode to the anode on the surface of the solid solution. The low-resistance metal link acts to short-out the relatively high-resistance glass and hence the overall resistance of the structure can be reduced by many orders of magnitude via this electrically stimulated deposition process. This is the basis of the Programmable Metallization Cell (PMC) technology [3]. In this paper, we will concentrate on results from the arsenic–trisulfide–silver system, doped by photo-dissolution to form a ternary composition near $Ag_4As_2S_3$. We will discuss the initial results of electrical characterization of two types of test structure: (1) devices in which a silver electrodeposit is formed on a $Ag_4As_2S_3$ channel between coplanar electrodes (Fig. 1A) and then subsequently 'broken' using a current pulse to form conducting islands separated by nanoscale gaps and (2) devices which have the solid solution deposited in a via in a dielectric which separates two metal layers, the electrodeposit being formed at the interface of the solid solution and the surrounding dielectric (Fig. 1B).

To form the coplanar devices, undoped As_2S_3, 40 nm thick, was evaporated from a powder source onto patterned resist on SiO_2 on a silicon substrate and immediately coated with a 25 nm thick layer of silver without breaking vacuum. The wafers were then removed and exposed to ultraviolet light at 436 nm for 10 min to fully saturate the chalcogenide and form the $Ag_4As_2S_3$ final composition (confirmed by RBS).
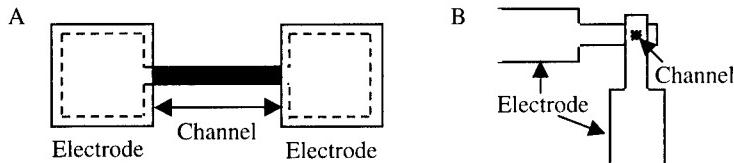
Fig. 1. A, Top view schematic of coplanar electrode device. B, Solid solution in via between electrodes.
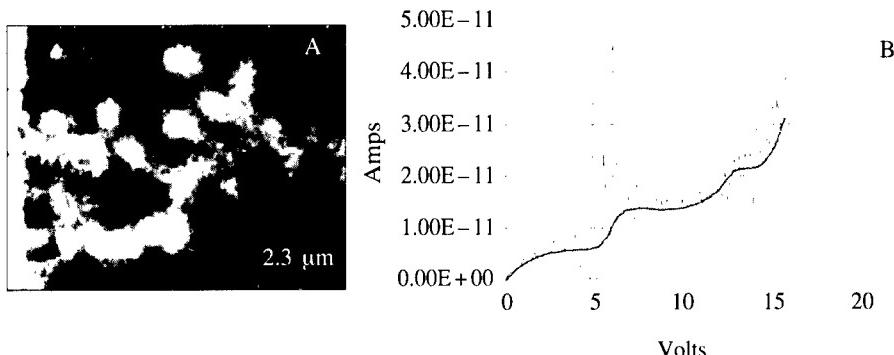

Fig. 2. A, Electron micrograph of dendritic electrodeposit which has been broken to form isolated silver islands (the anode is on the far left). B, $I-V$ plot of device at 4.2 K (the solid curve is a guide for the eye).

After removal of excess silver, the doped chalcogenide is then patterned by lift-off. The large silver contacts are also deposited by evaporation and patterned by lift-off. Using a device with 10 μm anode-to-cathode spacing, a narrow bridging dendritic electrodeposit was formed by applying 2 V across the electrodes with a compliance of 10 μA. This led to a room-temperature electrodeposit resistance of approximately 58 kΩ. The device was then cooled to 4.2 K by immersion in liquid He and the resistance decreased to around 3 kΩ, suggesting that the overall resistance is influenced by the grain boundaries in the electrodeposit. A 10 mA current pulse was then applied to break the electrodeposit by local thermal dissolution at the grain boundaries to form metallic silver islands separated by nanoscale gaps as shown in Fig. 2A. The current-voltage characteristic for this modified structure is given in Fig. 2B.

Room-temperature characterization of an 'open circuit' device such as this gives an exponential $I-V$ relationship in the GΩ range near $V = 0$, but the most obvious features of the plot of Fig. 2B are the steps at 5 V intervals in the $I-V$ characteristic, approximately 7 pA in height. This characteristic is similar to that obtained in devices which exhibit single electron modulated conduction in channels of nanocrystalline silicon [4, 5]. However, the complex nature of transport at low temperatures in these structures is revealed by the $I-V$ plots of Fig. 3. The curves of Fig. 3A, labeled 1–4, represent the $I-V$ characteristics of the device under the following conditions: (1) initial 0–10 V sweep, (2) subsequent 0–10 V sweep, (3) initial 0 to −10 V sweep, (4) subsequent 0 to −10 V sweep. It is clear from these curves that the $I-V$ characteristics of the structure depend on biasing history. This is reinforced by the results shown in Fig. 3B which are the $I-V$ curves for 0–10 V sweeps (only 0–5 V is shown for clarity) following exposure to increasing negative bias. The curve on the far right is the initial characteristic produced by repeated 0–10 V sweeps. The next curve to the left was produced after biasing the structure with −1 V, the next with −2 V, then −3 V, −5 V, and the leftmost curve is after −10 V.

The step-like features apparent in Fig. 2B are suggestive of a Coulomb blockade effect. If we assume these features do indeed result from single-electron tunneling via a dominant bottleneck [5], which we could
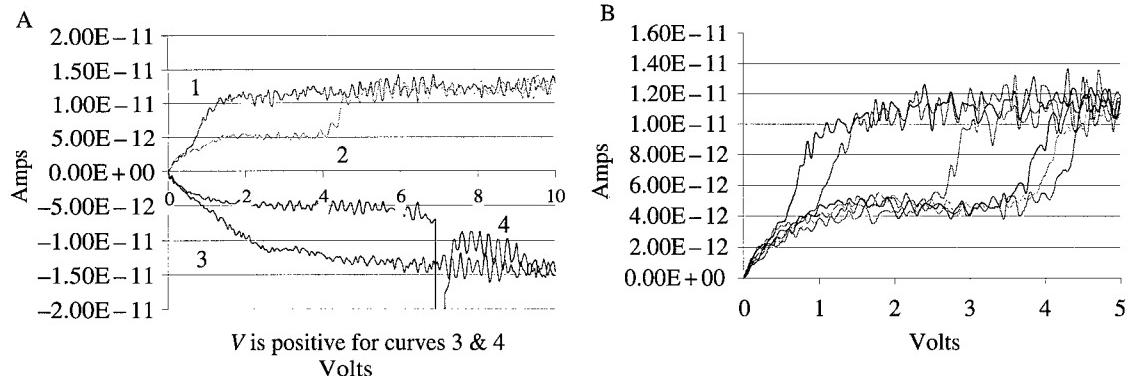
Fig. 3. A, I–V plots for initial and subsequent positive and negative biasing at 4.2 K. B, I–V plots after biasing with increasingly negative voltage ($+10$, -1, -2, -3, -5, -10 from right to left) at 4.2 K.

treat as an isolated Coulomb island with symmetric tunnel barriers, it is well known that current steps of height $e/2R_tC$ and width e/C should be expected in the current–voltage characteristic. Here, R_t is the resistance of the tunnel barriers and C is the total capacitance of the island. From the plateau width of approximately 5 V in Fig. 2B we thus obtain $C = 3.2 \times 10^{-20}$ F, while from the current steps of 7 pA we obtain $R_t = 3.6 \times 10^{11}$ Ω. The extremely small value of the capacitance suggests that the Coulomb blockade effect itself might be mediated by single ions in the solid solution between the metallic islands. Other experimental features may point against such an interpretation however. It is clear from the current–voltage characteristics in Figs 2B and 3, for example, that there is no evidence for a Coulomb gap near zero bias. This observation might be accounted for by the existence of a background charge in the system, perhaps that of the (charged) double layer at the metal–solution interface. The magnitude of this charge would need to be such that it lifted the Coulomb blockade close to zero bias. The capacitance of 3.2×10^{-20} F corresponds to a charging energy $e^2/2C$ of order 10 eV, which is (perhaps coincidentally) close to the ionization energy of silver, but the staircase effect is not evident at room temperature. Finally, the only reasonable explanation for the bias history-dependent behavior is that the bias alters the position of the ions in solution which in turn alters the nature of the assumed barrier in a continuous and reversible fashion. Since thermal diffusion is negligible at low temperature, the ions will remain where the field places them. Further studies are required to determine whether Coulomb effects are responsible for this interesting behavior.

Turning now to the fabrication of the vertical structures, shown schematically in Fig. 1B. The silver lower electrode was patterned using lift-off and covered with 40 nm of silicon nitride. Electron beam lithography, using a 100 nm thickness of 4% PMMA in chlorobenzene, was then used to define the vias in the nitride. The accelerating voltage was 40 kV and the dose was 4×10^{-8} C cm^{-1}. These conditions led to 50 nm wide vias following development and reactive ion etching of the nitride in a CF_4/O_2 plasma. The As_2S_3 layer, 25 nm thick, and 15 nm of silver was then deposited and the silver photodissolved prior to lift-off and PMMA removal in acetone. This left the solid solution in the via only. Figure 4 shows a 50 nm structure at this stage of the fabrication process. The top silver electrode was then deposited and patterned by lift-off. The off-resistance of this structure was in the order of 10^{12} Ω at 100 mV, which is actually several orders of magnitude higher than the solution resistance (due to the barrier-like characteristics of the metal–solid solution interface, i.e. the double layer). The devices switched to a low-resistance on-state by the formation of a bridging electrodeposit at around 1.8 V. The on-resistance was a strong function of the current compliance used but was as low as 100 Ω for 1 mA, illustrating that a 10 order of magnitude change in resistance is possible in the nanoscale switching element. The on-state proved to be extremely stable in these devices,

Fig. 4. Field emission scanning electron micrograph of solid solution in a 50 nm wide via in Si_3N_4 on an underlying silver electrode.

suggesting that they could have utility in extremely high-density programmable read-only memories (PROM) or antifuse elements.

In conclusion, we have seen that devices based on solutions of silver in arsenic trisulfide have interesting electrical characteristics at the nanoscale. Surface electrodeposits which have been broken using a current pulse exhibit voltage-programmable I–V characteristics at 4.2 K and permanent resistance reductions over 10 orders of magnitude are attainable in ultra-compact vertical devices.

References

[1] T. Kawaguchi, S. Maruno, and K. Tanaka, J. Non-Cryst. Solids **1231**, 164–166 (1993).

[2] Y. Hirose and H. Hirose, J. Appl. Phys. **47**, 2767 (1976).

[3] M. N. Kozicki and W. C. West, U. S. Patent Number 5,761,115 (1998).

[4] A. H. M. Kamal, J. Lutzen, B. A. Sanborn, M. V. Sidorov, M. N. Kozicki, D. J. Smith, and D. K. Ferry, Semicond. Sci. Technol. **13**, 1328 (1998).

[5] J. Lutzen, A. H. M. Kamal, M. N. Kozicki, D. K. Ferry, M. V. Sidorov, and D. J. Smith, J. Vac. Sci. Technol. **B16**, 2802 (1998).

The use of biopolymer templates to fabricate low-dimensional gold particle structures

C. A. Berven, M. N. Wybourne, L. Clarke[†]

Department of Physics and Astronomy, Dartmouth College, Hanover, NH 03755, U.S.A.

J. E. Hutchison, L. O. Brown, J. L. Mooster, M. E. Schmidt

Department of Chemistry, University of Oregon, Eugene, OR 97403, U.S.A.

We report the current–voltage characteristics of gold nanoparticle–biopolymer networks at room temperature. The characteristics have features that are indicative of single-electron charging in ordered, one-dimensional chains of nanoparticles. From capacitance estimates and numerical simulations, we argue that the observed electrical behavior is related to the low size dispersion of the nanoparticles and the uniformity of the biopolymer lengths imbedded within the network.

© 2000 Academic Press

Key words: nanoparticles, Coulomb blockade, biopolymer.

1. Introduction

Ligand-stabilized metal nanoparticles possess many qualities that facilitate the investigation of single-electron charging effects [1–5]. Because of their small size, typically <100 atoms, the Coulomb charging energy of the nanoparticles is at least an order of magnitude greater than the thermal energy ($k_B T$) at room temperature [6]. Nanoparticles of this size can support single-electron charging effects at 300 K when the ligand shell forms a tunnel barrier of resistance greater than the resistance quantum [7, 8]. In addition to providing electrical isolation, the ligand shell can be chemically modified to facilitate coupling between nanoparticles or attachment to other molecules [9].

We report transport measurements on gold nanoparticles with ligand shells attached to a rigid biopolymer network through the interaction between the ligand shell and the biopolymer. The room-temperature electrical conductance of all such samples shows a distinct Coulomb gap at low-bias voltages. Many samples showed periodic structure at voltages above threshold. Periodic conductance features are well known to be associated with single-electron effects [10–15]. What is unusual about the data reported here is the fact that the voltage scale of the periodicity is several orders of magnitude greater than that found in other quantum dot or mesoscopic samples.

[†]Present address: Department of Physics, University of Colorado at Boulder, Boulder, CO, U.S.A.
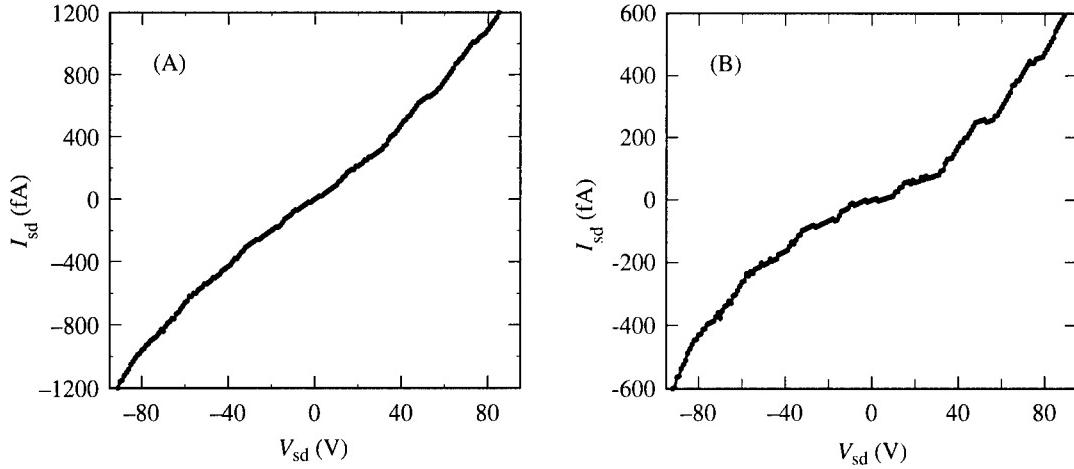
Fig. 1. A, The $I-V$ characteristic of a poly-L-lysine-nanoparticle sample. B, The $I-V$ data after subtraction of the background conductance of the undecorated biopolymer.

2. Experiment

Poly-L-lysine (54 000 amu) was deposited onto interdigitated gold electrodes with 15 μm spacings to form a framework to which gold nanoparticles were attached. The average length of the poly-L-lysine was 30 nm. The details of the fabrication have been reported elsewhere [16]. TEM measurements showed that the radius of the metal core was 0.7 nm, with a variation of $\pm 20\%$. The radius of the core and ligand shell together is estimated to be 2.1 nm.

Current–voltage ($I-V$) measurements were made in an electrically shielded vacuum chamber at room temperature [3]. Radio frequency (RF) electric fields could be applied by means of a dipole antenna placed close to the sample.

3. Results and discussion

Representative $I-V$ behavior of the biopolymer–nanoparticle samples is shown in Fig. 1A. All samples exhibited nonlinear behavior, with many showing structure equally spaced in voltage. Control experiments showed little or no difference between the $I-V$ characteristics of the undecorated biopolymer and the bare electrodes. When the data are corrected for the conductance of the undecorated poly-L-lysine, a blockade region and plateaus regularly spaced in voltage are clearly visible, as shown in Fig. 1B. For this sample the threshold voltage was $V_T = 12 \pm 1$ V, and the period of the oscillations was $\Delta V = 25 \pm 3$ V. In general, the ratio $\Delta V / V_T$ was frequently close to 2. Applying RF signals to the sample had no observable effect on the plateau structure.

Above threshold, the scaling $I \propto (V/V_T - 1)^\gamma$ was found to describe all sets of data, with $\gamma = 1.2 \pm 0.2$, as illustrated in Fig. 2. Here the error includes the uncertainty in the current measurement and the spread between different samples. Unlike the quadratic dependence found in thin films containing gold nanoparticles [3, 7, 8], this result is consistent with single-electron transport in both ordered and disordered one-dimensional systems where it is predicted that $\gamma \sim 1$ [17, 18]. Experiments on one-dimensional chains of tunnel junctions found $\gamma = 1.36 \pm 0.1$ [19]. Thus, from the current–voltage scaling above threshold, we conclude that regions of one-dimensional morphology dominate the transport in the biopolymer–nanoparticle network.

From the sample layout, the distance between the nanoparticle cores and the ground plane is expected
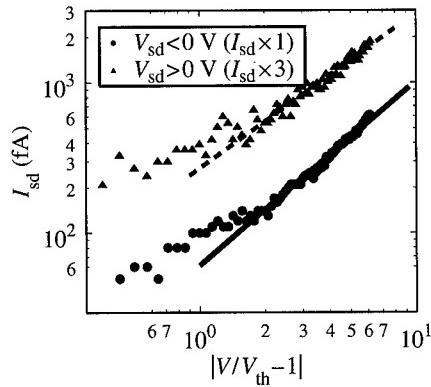
Fig. 2. Current versus scaled voltage. The solid circles (triangles) represent negative (positive) bias voltages. For clarity, the current values for positive bias have been multiplied by 3. The solid and dash lined are least square fits to the data and have slopes of 1.2 and 1.0, respectively.

to be relatively large. Thus the capacitance to ground, C_g, can be approximated by the self capacitance of a conducting sphere embedded in a dielectric shell. Assuming the ligand has a dielectric constant of 3, we estimate that $C_g \approx 0.14$ aF. Treating the adjacent nanoparticles as identical metal spheres of radius 0.7 nm whose centers are separated by a distance 4.2 nm, we estimate the maximum inter-particle capacitance to be $C_{dd} = 0.023$ aF. Using the above capacitance estimates in a conventional Coulomb blockade model [18] we obtain $V_T = 0.5$ V, which is at least an order of magnitude smaller than the threshold voltages measured.

Monte Carlo simulations [20] of the conductance characteristics of chains consisting of one, two and five nanoparticles with $C_g = 0.14$ aF and $C_{dd} = 0.023$ aF are shown in Fig. 3A. In the regime where $C_g \gg C_{dd}$, plateaus occur with a periodicity of about e/C_g. Also, these simulations show that the ratio $\Delta V/V_T$ is equal to 2 when there is little or no disorder and that the number of plateaus in the $I-V$ characteristics is approximately equal to the number of nanoparticles in the chain. The effect of a variation in the values of C_g and C_{dd} is shown in Fig. 3B. The capacitance disorder due to the uncertainty in the measured nanoparticle radius ($\pm 20\%$) is not found to affect the periodicity of the plateaus. However, introducing an uncertainty of $\pm 50\%$ in C_g ($\pm 67\%$ in radius) can destroy the periodicity. Variation of C_{dd} by $\pm 50\%$ had little or no effect on the $I-V$ characteristics.

The simulations demonstrate that chains of nanoparticles where $C_g \gg C_{dd}$ will have $I-V$ characteristics with properties similar to the data. However, the voltage scale needs to be addressed. Although conventional Coulomb staircase phenomena have been reported at 287 K in long chains of particles [12], to reproduce the structure seen in our data from a simple chain of clusters would require the nanoparticles to be unphysically small. For this reason, the origin of the structure is unlikely to come from a few smaller nanoparticles dispersed in the chain.

The voltage scale discrepancy may be explained if the biopolymer–nanoparticle network contains sections that are not in Coulomb blockade. These sections could form a bias network for those in blockade, thus reducing the voltage across each section in blockade. The weakly conducting sections could result from regions in which the capacitance of the nanoparticles is increased. Also, the background conductance between the electrodes themselves could provide a bias network. The size of the electrodes causes a high degree of spatial averaging over the network. Therefore, if the idea of a resistive bias network is correct it suggests that the conductance oscillations are insensitive to the local arrangement of the poly-L-lysine network. This implies that their origin must be related to system parameters that have a high degree of uniformity, such as the molecular weight (length) of the poly-L-lysine, or the size (capacitance) of the nanoparticles.
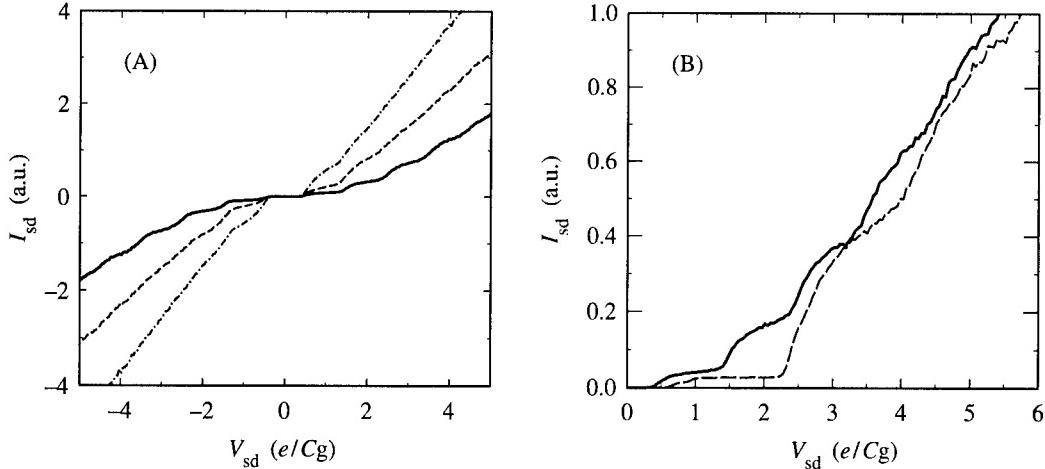
Fig. 3. A, Monte Carlo simulations of chains of one (— • —), two (— —) and five (——) nanoclusters in series. B, Monte Carlo simulations of five nanoclusters in series with ±20% (——) and ±67% (— —) variation in nanocluster radius. All simulations were done for a temperature $T = 300$ K.

4. Summary

In summary, we have measured the room-temperature transport in a biopolymer–nanoparticle system. The almost linear relationship between current and voltage above threshold suggests that the poly-L-lysine molecules act as a one-dimensional template for the gold nanoparticles. From the observed plateau structure and capacitance estimates, we believe the one-dimensional sections of the network contain at least five nanoparticles. Single-electron charging effects are expected to be sensitive to the size distribution of nanoparticles and to the length of the poly-L-lysine biopolymer, which may be the reason we do not observe the plateau structure in all samples.

References

[1] G. Schön and U. Simon, Colloid. Polym. Sci. **273**, 202 (1995).

[2] D. L. Feldheim and C. D. Keating, Chem. Soc. Rev. **27**, 1 (1998).

[3] L. Clarke, M. N. Wybourne, M. Yan, S. X. Cai, and J. F. W. Keana, Appl. Phys. Lett. **71**, 617 (1997).

[4] R. P. Andres, J. D. Beielefeld, J. I. Henderson, D. B. Janes, V. R. Kolagunta, C. P. Kubiak, W. J. Mahoney, and R. G. Osifchin, Science **273**, 1690 (1996).

[5] T. Sato, H. Ahmed, D. Brown, and B. F. G. Johnson, J. Appl. Phys. **82**, 696 (1997).

[6] R. P. Andres, T. Bein, M. Dorogi, S. Feng, J. I. Henderson, C. P. Kubiak, W. Mahoney, R. G. Osifchin, and R. Reifenberger, Science **272**, 1323 (1996).

[7] L. Clarke, M. N. Wybourne, M. Yan, S. X. Cai, L. O. Brown, J. Hutchison, and J. F. W. Keana, J. Vac. Sci. Technol. **B15**, 2925 (1997).

[8] M. N. Wybourne, J. E. Hutchison, L. Clarke, L. O. Brown, and J. L. Mooster, Microelec. Eng. **47**, 55 (1999).

[9] L. O. Brown and J. E. Hutchison, J. Am. Chem. Soc. **119**, 12384 (1997).

[10] D. V. Averin and K. K. Likharev, *Mesoscopic Phenomena in Solids*, edited by B. Al'tshuler, P. Lee, and R. A. Webb (Elsevier, Amsterdam, 1991).

[11] N. S. Bakhvolov, G. S. Kazacha, K. K. Likharev, and S. I. Serdyukova, Sov. Phys. JETP **68**, 581 (1989).

[12] K. Kawaski, M. Mochizuki, and K. Tsutsui, Jpn. J. Appl. Phys. **38**, 418 (1999).

[13] L. Y. Gorelik, A. Isacsson, M. V. Voinova, B. Kasemo, R. I. Shekhter, and M. Jonson, Phys. Rev. Lett. **80**, 4526 (1998); A. Isacsson, L. Y. Gorelik, M. V. Voinova, B. Kasemo, R. I. Shekter, and M. Jonson, Physica **B255**, 150 (1998).

[14] M. Shin, S. Lee, K. W. Park, and E. Lee, Phys. Rev. Lett. **80**, 5774 (1998).

[15] D. C. Ralph, C. T. Black, and M. Tinkham, Phys. Rev. Lett. **74**, 3241 (1995); M. Tinkham, Am. J. Phys. **64**, 343 (1996).

[16] M. N. Wybourne, L. Clarke, C. A. Berven, J. E. Hutchison, L. O. Brown, and J. L. Mooster, MRS proceedings in press.

[17] A. A. Middleton and N. S. Wingreen, Phys. Rev. Lett. **71**, 3198 (1993).

[18] G. Y. Hu and R. F. O'Connell, Phys. Rev. **B49**, 16773 (1994).

[19] A. J. Rimberg, T. R. Ho, and J. Clarke, Phys. Rev. Lett. **74**, 4714 (1995).

[20] SIMON–Simulation of Nano Structures–Christoph Wasshuber.

Multistability, absolute negative conductivity and spontaneous current generation in semiconductor superlattices in large magnetic fields

ETHAN H. CANNON[†], DAVID K. CAMPBELL

*Department of Physics, University of Illinois at Urbana-Champaign, 1110 W. Green St., Urbana,
IL 61801, U.S.A.*

FEODOR V. KUSMARTSEV

Department of Physics, Loughborough University, Loughborough LE11 3TU, U.K.

KIRILL N. ALEKSEEV

Theory of Nonlinear Processes Laboratory, Kirensky Institute of Physics, Krasnoyarsk 660036, Russia

We discuss electron transport through a semiconductor superlattice subject to an electric field parallel to, and a magnetic field perpendicular to, the growth axis using a semiclassical balance equation model. We find that the current–voltage characteristic becomes multistable in a large magnetic field; furthermore, hot electrons display novel features in their current–voltage characteristic, including absolute negative conductivity and a spontaneously generated dc current at zero bias.

© 2000 Academic Press

Key words: superlattice, miniband transport, nonlinear transport.

Semiconductor superlattices (SSLs) offer the exciting prospect of experimentally observing nonlinear transport and optical properties. The large period along the growth axis leads to narrow minibands; consequently, moderate electric and magnetic fields can accelerate electrons into the nonparabolic regions of the minibands before a scattering event occurs. Electrons which explore a significant portion of the miniband exhibit such nonlinear properties as negative differential conductivity (NDC) and Bloch oscillations; for an introduction, see [1].

We study a balance equation model for transport through a single miniband SSL with electric field, E, along the growth axis (the z-direction) and magnetic field, B, in the plane of the quantum wells (QWs) that form the SSL (the x-direction). These balance equations are derived from the semiclassical Boltzmann transport equation (BTE) which describes miniband transport in electric or magnetic fields not strong enough to localize electrons within a single period of the SSL. The balance equations for an electric field were first derived in [2], while the extension to include the magnetic field was discussed in [3,4]; and we refer the reader to these references for the detailed derivation.

We assume a tight-binding miniband dispersion relation, $\epsilon(k) = \hbar^2 k_y^2 / 2m^* + \Delta/2[1 - \cos(k_z a)]$, where

---

m^* is the effective mass in the plane of the QWs, Δ is the miniband width, and a is the SSL period. The balance equations are

$$\dot{V}_y = -\frac{eB}{m^*c}V_z - \gamma_{vy}V_y \tag{1}$$

$$\dot{V}_z = -\frac{e}{m(\varepsilon_z)}\left[E - \frac{BV_y}{c}\right] - \gamma_{vz}V_z \tag{2}$$

$$\dot{\varepsilon}_z = -eEV_z + \frac{eB}{c}V_yV_z - \gamma_\varepsilon[\varepsilon_z - \varepsilon_{eq,z}], \tag{3}$$

where $-e$ is the charge of an electron and c is the speed of light. The average electron velocity, $V = (V_y, V_z)$, is obtained by integrating the distribution function which satisfies the BTE over the Brillouin zone; γ_{vy} and γ_{vz} are the damping rates for these quantities, which follow from inelastic phonon scattering and elastic impurity and interface roughness scattering. Likewise, ε_z represents the average energy of motion along the growth axis; $\varepsilon_{eq,z}$ is its equilibrium value, and γ_ε is its relaxation rate due mainly to inelastic scattering. The balance equations contain the effective mass term $m(\varepsilon_z) = m_0/(1 - 2\varepsilon_z/\Delta)$, where $m_0 = 2\hbar^2/\Delta a^2$ is the effective mass at the bottom of the miniband. Owing to the nonparabolic dispersion relation, the effective mass for motion along the growth axis depends on the corresponding energy component; in contrast, the energy of motion within the QW does not enter the balance equations since the effective mass for this motion is constant. While the magnetic field does not change the total electron energy, it does transfer energy between the longitudinal and transverse directions, hence it affects the time dependence of ε_z.

The following scalings facilitate numerical studies: $v_y = ((m_0m^*)^{1/2}a/\hbar)V_y$, $v_z = (m_0a/\hbar)V_z$, $w = (\varepsilon_z - \Delta/2)/(\Delta/2)$, $w_0 = (\varepsilon_{eq,z} - \Delta/2)/(\Delta/2)$, $\mathcal{B} = eB/(m_0m^*)^{1/2}c$ and $\omega_B = eEa/\hbar$ (the Bloch frequency of the electric field). Accordingly, the set of balance equations becomes

$$\dot{v}_y = -\mathcal{B}v_z - \gamma_{vy}v_y \tag{4}$$

$$\dot{z}_z = \omega_B w - \mathcal{B}v_y w - \gamma_{vz}v_z \tag{5}$$

$$\dot{w} = -\omega_B v_z + \mathcal{B}v_y v_z - \gamma_\varepsilon(w - w_0). \tag{6}$$

With the time-independent electric field ω_B and magnetic field $\mathcal{B}$, the SSL current $I = -eNA(\Delta a/2\hbar)v_{z,ss}$, where N is the carrier concentration, A is the cross-sectional area, and $v_{z,ss}$ is the steady-state solution to eqn (5). Considering the steady-state solutions to eqns (4)–(6), we obtain a cubic equation relating $v_{z,ss}$ to the applied voltage, with $C = \mathcal{B}^2/\gamma_{vy}$

$$C^2 v_{z,ss}^3 + 2C\omega_B v_{z,ss}^2 + [\gamma_{vz}\gamma_\varepsilon + \omega_B^2 - \gamma_\varepsilon w_0 C]v_{z,ss} - \gamma_\varepsilon w_0 \omega_B = 0. \tag{7}$$

The current–voltage characteristic of an SSL exhibits a peak followed by NDC; and a magnetic field in the plane of the QWs increases the critical electric field at which the peak current is attained [5,6]. As is evident in Fig. 1, the balance equations reproduce these well-known results; moreover, they predict the new result of multistability for a sufficiently large magnetic field (Fig. 1C). For the very low relaxation rates obtained in very recent experiments—$\gamma_{vy} = \gamma_{vz} = \gamma_\varepsilon = 1.5 \times 10^{13}$ s^{-1} for an SSL with $\Delta = 23$ meV and $a = 84$ Å [7]—this multistability may be observable at a modest magnetic field of a couple of Tesla.

The average energy for motion along the growth axis relaxes to its equilibrium value w_0 at the rate γ_ε; for thermal carriers, $w_0 \leq 0$. We now wish to discuss the situation where $w_0 > 0$; this occurs for 'hot' electrons which have a highly nonequilibrium distribution, even without applied fields, and requires a constant influx of energy. Experimentally, these hot electrons can be obtained by injecting electrons into the top half of the miniband of a finite SSL, as in the recent experiment described in [7]. Consider the situation in which there are no external fields: if electrons leave the SSL before relaxing to the bottom of the miniband, and new electrons replace them at the injection energy in the top half, then $w = w_0 > 0$. The energy to maintain $w_0 > 0$ comes from the steady injection of energetic electrons.
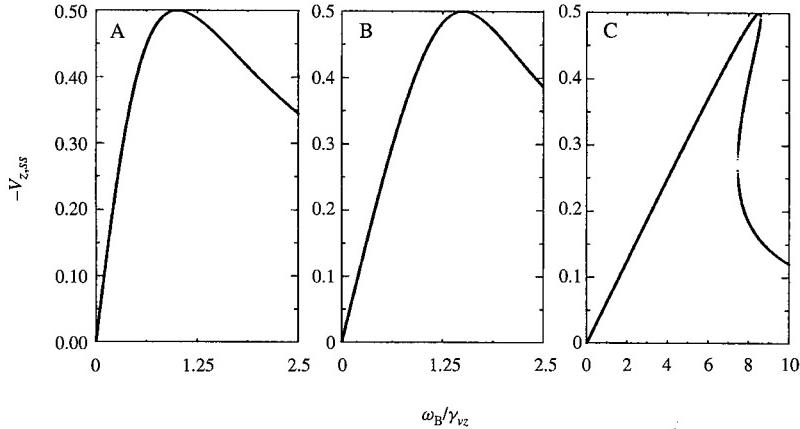
Fig. 1. Scaled current–voltage characteristic for an SSL with $w_0(\gamma_\varepsilon/\gamma_{vz})^{1/2} = -1$ and $C = 0$ (A), $C = \gamma_{vz}$ (B) and $C = 15\gamma_{vz}$ (C).
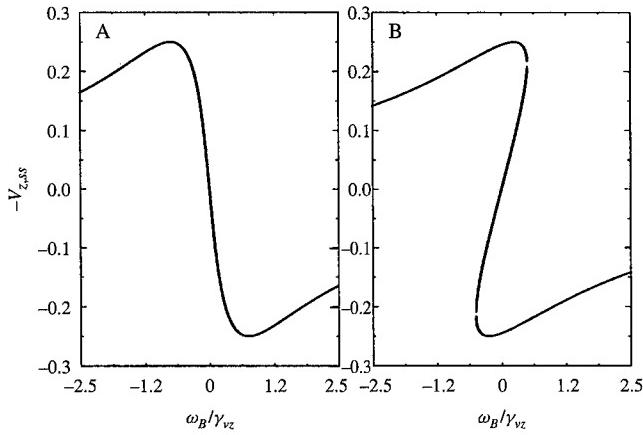

Fig. 2. Scaled current–voltage characteristic for hot electrons in an SSL with $w_0(\gamma_\varepsilon/\gamma_{vz})^{1/2} = 0.5$ and $C = \gamma_{vz}$ (A), and $C = 5\gamma_{vz}$ (B).

The balance equations predict two novel features in the current–voltage characteristic for hot electrons: absolute negative conductivity (ANC) and, for sufficiently large magnetic fields, spontaneous current generation at zero bias. Figure 2A illustrates the ANC, as a positive bias induces a negative current when $w_0 > 0$. For the larger magnetic field in Fig. 2B, multistability occurs at zero-bias; in fact, the zero bias–zero current solution is unstable and the SSL spontaneously developed a current across it. For ANC, the energy of the hot electrons supports a current against the applied bias; similarly, spontaneous current generation occurs when the hot electrons use their energy to maintain a current in the absence of any bias. Both effects may be observable through their influence on the current–voltage characteristics of high-quality samples [3].

References

[1] *Semiconductor Superlattices: Growth and Electronic Properties*, edited by H. T. Grahn (World Scientific, New Jersey, 1995).

[2] A. A. Ignatov and V. I. Shashkin, A simplified approach to nonlinear hf response theory of superlattice materials, Phys. Lett. **A94**, 169 (1983).

[3] E. H. Cannon, F. V. Kusmartsev, K. N. Alekseev, and D. K. Campbell, Absolute negative conductivity and spontaneous current generation in semiconductor superlattices with hot electrons, cond-mat/9910042.

[4] E. H. Cannon, Nonlinear transport in semiconductor superlattices, PhD thesis, University of Illinois at Urbana-Champaign, (1999).

[5] F. Aristone, J. F. Palmier, P. Gassot, J. C. Portal, and F. Mollot, Transverse magnetotransport anisotropy in a semiconductor superlattice, Appl. Phys. Lett. **67**, 2916 (1995).

[6] L. Canali, F. Beltram, M. Lazzarino, and L. Sorba, High-field transport in superlattices: Observation of the Stark-cyclotron resonance, Superlatt. Microstruct. **22**, 155 (1997).

[7] C. Rauch, G. Strasser, K. Unterrainer, W. Boxleitner, E. Gornik, and A. Wacker, Transition between coherent and incoherent electron transport in GaAs/GaAlAs superlattices, Phys. Rev. Lett. **81**, 3495 (1998).

Effect of boundaries and impurities on electron–phonon dephasing

A. SERGEEV, V. MITIN

Department of ECE, Wayne State University, Detroit, MI 48202, U.S.A.

Electron scattering from boundaries and impurities destroys the single-particle picture of the electron–phonon interaction. We show that quantum interference between 'pure' electron–phonon and electron–boundary/impurity scattering may result in the reduction as well as to the significant enlargement of the electron dephasing rate. This effect crucially depends on the extent, to which electron scatterers, such as boundaries and impurities, are dragged by phonons. Static and vibrating scatterers are described by two dimensionless parameters $q_T l$ and $q_T L$, where q is the wavevector of the thermal phonon, l is the total electron mean-free path, L is the mean-free path due to scattering from static scatterers. According to the Pippard ineffectiveness condition [1], without static scatterers the dephasing rate at low temperatures is slower by the factor $1/ql$ than the rate in a pure bulk material. However, in the presence of static potential the dephasing rate turns out to be $1/qL$ times faster. Thus, at low temperatures electron dephasing and energy relaxation may be controlled by electron boundary/impurity scattering in a wide range.

© 2000 Academic Press

Key words: electron–phonon interaction, dephasing, disorder.

1. Introduction

In recent years, the electron dephasing in ultrathin films, nanostructures and mesoscopic devices has been intensively studied. Temperature-dependent dephasing rate is mainly determined by the electron–electron and electron–phonon interactions. While theoretical results pertaining to electron–electron scattering are confirmed by many experiments, the electron–phonon mechanism is still poorly understood. Unfortunately, most researchers employ the standard clean-limit concept, its uncritical application leads to incorrect and controversial conclusions. A reliable electron–phonon interaction model taking into account electron scattering from boundaries, defects and impurities is of vital importance.

If scattering potential of boundaries and impurities is completely dragged by phonons, the inelastic electron scattering from this potential may be excluded by a transformation to the frame, which moves together with the phonon. Using transformation to the local frame, Pippard [1] has found that the electron–phonon coupling depends substantially on the parameter ql, where q is the wavevector of the phonon, and l is the electron mean-free path. If $ql < 1$, the electron–phonon coupling is a factor of $1/ql$ weaker than the coupling in the pure limit, $l \rightarrow \infty$. This statement is well known as the Pippard ineffectiveness condition [2]. It was confirmed by microscopic calculations in [3, 4].

Obviously the Pippard's assumption about completely dragged scatterers is not valid in the presence of rigid boundaries or heavy defects. To generalize Pippard's model we take into account additional static potential. We show that even relatively weak static potential drastically changes the effective electron–phonon coupling and corresponding electron dephasing rate.

2. Model

We start with the Hamiltonian, which describes the 'pure' electron–phonon interaction and the interaction between electrons and scatterers that are completely dragged by phonons [3],

$$
H_{int} = \sum_{p,q} g(q) c_{p+q}^+ c_p (b_{q,n} + b_{-q,n}^+) + \sum_{p,k,R_\alpha} V(k) c_p^+ c_{p-k} \exp(-ikR_\alpha)
$$
$$
+ \sum_{p,k,q,R_\alpha} \gamma(k,q) c_p^+ c_{p-k} (b_{q,n} + b_{-q,n}^+) \exp[-i(k-q)R_\alpha], \tag{1}
$$

where c_p^+ is the electron creation operator, $b_{q,n}^+$ is the creation operator of a phonon with a wavevector q and polarization index n, and R_α are the equilibrium positions of scatterers.

The first term with the vertex $g = (2\epsilon_F/3) \times (q \cdot e_n)/(2\rho\omega)^{1/2}$ (ϵ_F is the Fermi energy, e_n is the phonon polarization vector, ρ is the density) corresponds to the pure electron–phonon scattering. The second term describes the elastic electron scattering from the potential of boundaries and impurities, $V(k)$. If this potential is completely dragged by phonons, the vertex of inelastic electron scattering is given by [3,5]: $\gamma(k,q) = -iV(ke_n)/(2\rho\omega_q)^{1/2}$.

Now we take into account the static scatterers, such as rigid boundaries and heavy defects. Then the total momentum relaxation rate is $\tau^{-1} = \tau_d^{-1} + \tau_s^{-1}$, where τ_d^{-1} is the electron momentum relaxation rate due to scatterers that are dragged by phonons, and τ_s^{-1} is the relaxation rate due to static scatterers. It is convenient to introduce the electron mean-free path, $l = v_F\tau$, and the electron-free path with respect to scattering from static potential, $L = v_F\tau_s$.

3. Electron dephasing rate

Calculations employing the Keldysh diagrammatic technique for nonequilibrium processes show that the collision integral, which describes the interaction between longitudinal phonons and electrons in a disordered conductor with static and vibrating scattering potentials, is given by

$$
I_{e-l.ph}(\epsilon) = -\frac{2\beta_l}{(p_F u_l)^2} \int d\omega_q \, \omega_q^2 R(\epsilon, \omega_q) \left[\frac{ql \arctan(ql)}{ql - \arctan(ql)} - \left(1 - \frac{l}{L}\right) \frac{3}{(ql)} \right], \tag{2}
$$

where $\zeta_0 = \arctan(ql)/(ql)$, the dimensionless constant of the electron–phonon interaction $\beta_l = (2\epsilon_F/3)^2 \nu/2\rho u_l^2$ (ν is the electron density of states, u_l is the longitudinal sound velocity), and $R(\epsilon, \omega_q)$ is the combination of electron (n_ϵ) and phonon (N_ω) distribution functions: $R(\epsilon, \omega) = N_\omega n_\epsilon (1 - n_{\epsilon+\omega}) - (1 + N_\omega)(1 - n_\epsilon) n_{\epsilon+\omega}$. Then the dephasing/relaxation rate of electrons is given by

$$
\frac{1}{\tau_{e-l.ph}(0)} = \frac{7\pi\zeta(3)}{2} \frac{\beta_l T^3}{(p_F u_l)^2} F_l(q_T l), \tag{3}
$$

$$
F_l(z) = \frac{2}{7\zeta(3)} \int_0^{A_l} dx \, \Phi_l(xz) (N_x + n_x) x^2, \quad \Phi_l(x) = \frac{2}{\pi} \left[\frac{x \arctan(x)}{x - \arctan(x)} - \left(1 - \frac{l}{L}\right) \frac{3}{x} \right], \tag{4}
$$

where $A_{t(l)} = \theta_D l / u_{t(l)} z$ (θ_D is the Debye temperature). In the limiting cases the relaxation rate is

$$
\frac{1}{\tau_{e-l.ph}(0)} = \frac{7\pi\zeta(3)}{2} \frac{\beta_l T^3}{(p_F u_l)^2} \times \begin{cases} 1, & Tl > u_l, \\ \frac{2\pi^3}{35\zeta(3)} \frac{Tl}{u_l} + \frac{3\pi}{7\zeta(3)} \frac{u_l}{TL}, & Tl < u_l. \end{cases} \tag{5}
$$
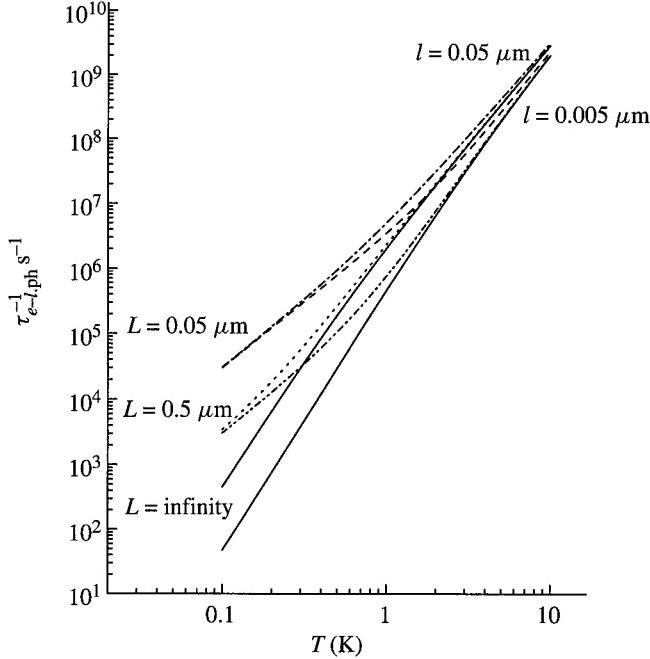
Fig. 1. Electron dephasing rate due to longitudinal phonons in Al structures with electron mean-free path $l = 0.005$ μm and 0.05 μm. Solid lines correspond to complete drag of all scatterers (boundaries and impurities) by phonons. Dashed and dotted lines correspond to the electron mean-free path with respect to the static potential $L = 0.05$ μm and 0.5 μm.

Now we consider interaction of electrons and transverse phonons. The corresponding collision integral has a form

$$I_{e-t.ph}(\epsilon) = -\frac{12\beta_t T^2}{(p_F u_t)(p_F l)}\left(1 - \frac{l}{L}\right)\int d\omega_q R(\epsilon, \omega_q)\left[1 + \left(1 - \frac{l}{L}\right)\frac{3ql - 3((ql)^2 + 1)\arctan(ql)}{2(ql)^3}\right], \quad (6)$$

where the dimensionless constant is $\beta_t = \beta_l(u_l/u_t)^2$, and u_t is the transverse sound velocity. Then the electron dephasing/relaxation rate is given by

$$\frac{1}{\tau_{e-t.ph}(0)} = \frac{3\pi^2\beta_t T^2}{p_F^2 u_t}\left(\frac{1}{l} - \frac{1}{L}\right)F_t(q_T l), \quad (7)$$

$$F_t(z) = \frac{4}{\pi^2}\int_0^{A_t} dx\,\Phi_t(xz)(N_x + n_x^{eq})x, \quad \Phi_t(x)$$

$$= 1 + \left(1 - \frac{l}{L}\right)\frac{3x - 3(x^2 + 1)\arctan(x)}{2x^3}. \quad (8)$$

In the limiting cases the electron relaxation rate is

$$\frac{1}{\tau_{e-t.ph}(0)} = \frac{3\pi^2\beta_t T^2}{p_F^2 u_t}\left(\frac{1}{l} - \frac{1}{L}\right) \times \begin{cases} 1, & Tl > u_t, \\ \frac{l}{L} + \left(1 - \frac{l}{L}\right)\frac{\pi^2}{10}\left(\frac{Tl}{u_t}\right)^2, & Tl < u_t. \end{cases} \quad (9)$$

Therefore, in the impure case ($Tl < u_l, u_t$), the total electron dephasing/relaxation rate is given by

$$\frac{1}{\tau_{e-ph}(0)} = \frac{\pi^4 T^4}{5}(p_F l)\left[\frac{\beta_l}{(p_F u_l)^3} + \frac{3\beta_t}{2(p_F u_t)^3}\left(1 - \frac{l}{L}\right)\right] + \frac{3\pi^2 T^2}{2p_F L}\left[\frac{\beta_l}{p_F u_l} + \frac{2\beta_t}{p_F u_t}\left(1 - \frac{l}{L}\right)\right]. \quad (10)$$
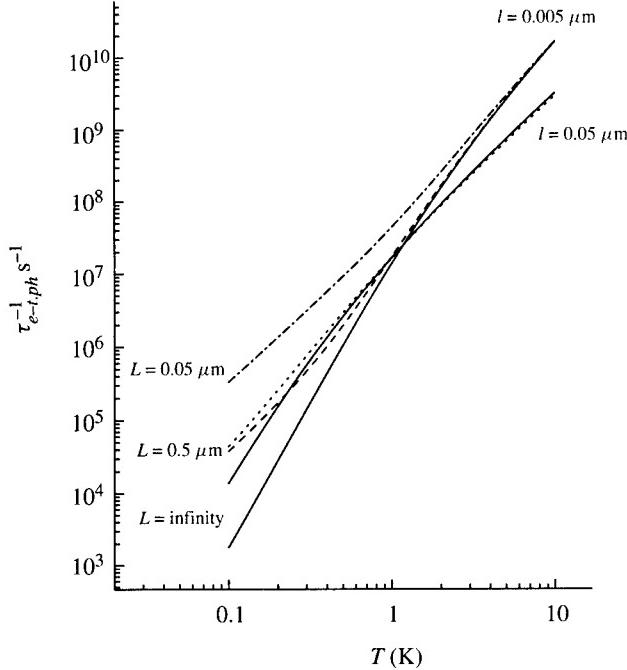
Fig. 2. Electron dephasing rate due to transverse phonons in Al structures with electron mean-free path $l = 0.005$ μm and 0.05 μm. Solid lines correspond to complete drag of all scatterers (boundaries and impurities) by phonons. Dashed and dotted lines correspond to the electron mean-free path with respect to the static potential $L = 0.05$ μm and 0.5 μm.

4. Discussion

The presence of static and vibrating electron scatterers leads to complex quantum interference between different scattering mechanisms. If boundaries, defects and impurities are completely dragged by phonons ($L \to \infty$), we reproduce results of Ref. [4] for the electron–phonon dephasing rate. In agreement with the Pippard ineffectiveness condition, at low temperatures the electron–phonon dephasing rate is $u/(Tl)$ times slower than the rate in a pure conductor, $1/\tau_{e-ph} \propto T^4 l$. Note, that in our model with spherical Fermi surface, only longitudinal phonons interact with electrons in the pure conductor. Inelastic electron scattering from boundaries and impurities generates a new channel of the electron–phonon interaction. Due to vibrating boundaries and impurities, transverse phonons can interact with electrons. In the pure limit, $Tl/u \gg 1$, this channel is Tl/u times weaker than the pure electron–phonon coupling. In the dirty limit, $Tl/u \ll 1$, both channels are enhanced due to the diffusion electron motion, in the same way as the electron–electron interaction in impure conductors. However, the quantum interference between these two channels of electron scattering neglects the total effects and results in the Pippard ineffectiveness condition.

This picture is changed in the presence of additional static potential or due to incomplete drag of boundaries and impurities by phonons. In the limit $Tl/u \ll 1$, where the interference is important, the electron dephasing rate turns out to be $u/(TL)$ times faster than the rate in the pure conductor, $1/\tau_{e-ph} \propto T^2/L$. Note, that compared with the dephasing rate due to longitudinal phonons, the contribution of transverse phonons consists of the large factor $(u_l/u_t)^3$ and the factor $(1 - l/L)$. The factor $(1 - l/L)$ has a simple interpretation: it is proportional to the concentration of vibrating scatterers, which provide the interaction between electrons and transverse phonons. If this factor is not too small, the effect of transverse phonons dominates at low temperatures. Note, that the T^2/l-term in the relaxation rate due to vibrating impurities
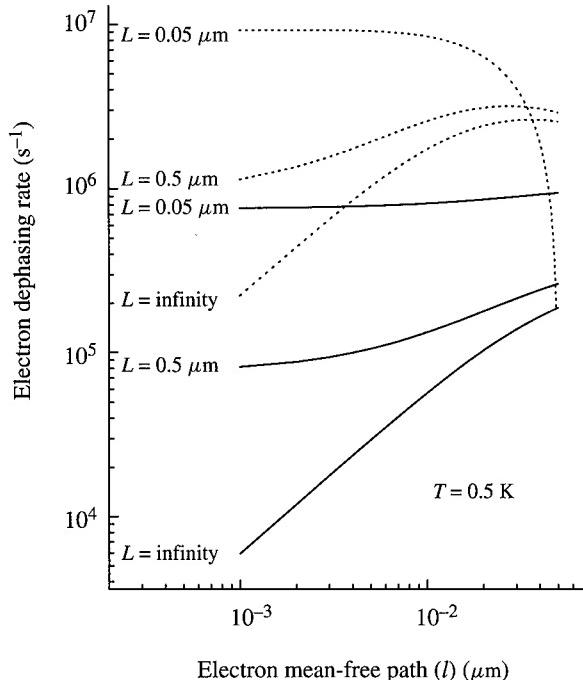
Fig. 3. Dependence of the dephasing rate on the electron mean-free path. Solid and dotted lines represent contributions of longitudinal and transverse phonons, correspondingly.

has been obtained by many authors [6]. This result is wrong and contradicts to the Pippard ineffectiveness concept. Only in the presence of the static potential do we obtain a T^2-term proportional to $1/L$.

To illustrate our results, we calculate the electron dephasing rate in Al structures with $u_l = 6.3 \times 10^5$ cm s^{-1}, $u_t = 3.1 \times 10^5$ cm s^{-1}, $v_F = 1.3 \times 10^8$ cm s^{-1}, $\beta_l = 1.14$, and $\beta_t = 4.7$ [5]. Temperature dependencies of the dephasing rate in the structures with the electron mean-free path 0.05 μm and 0.005 μm are presented in Figs 1 and 2. Solid lines show the dephasing rate under the Pippard ineffectiveness condition. At low temperatures the dephasing becomes faster in the presence of the static potential. Comparing Figs 1 and 2, we see that at low temperatures the electron dephasing is determined by transverse phonons. Figure 3 shows the dependence of the dephasing rate on the electron mean-free path. In the case of complete drag of boundaries and defects, the dephasing rate is proportional to l at low temperatures. In the presence of the static potential, the relaxation rate is determined mainly by the electron mean-free path with respect to scattering from the static potential.

It is important for applications, that the electron–phonon dephasing rate in mesoscopic devices may be changed in a wide range. It can be increased or decreased compared to the rate in a pure bulk material. Some experimental data support our conclusions. The enhancement of the electron–phonon interaction due to disorder has been found in thin metallic films [7] and semiconducting heterostructures [8]. The T^2-dependence of the electron–phonon dephasing rate is widely observed in experiments [9]. Some important points, such as the modification of the phonon spectrum, deserve further theoretical investigations.

References

[1] A. B. Pippard, Phil. Mag. **46**, 1104 (1955).

[2] J. M. Ziman, *Electrons and Phonons* (Clarendon, Oxford, 1960) p. 213.

[3] G. Grünvald and K. Scharnberg, Z. Phys. **268**, 197 (1974).

[4] J. Rammer and A. Schmid, Phys. Rev. **B34**, 1352 (1986).

[5] M. Yu. Reizer and A. V. Sergeev, Zh. Exsp. Teor. Fiz. **92**, 2291 (1987); [Sov. Phys. JETP **65**, 1291 (1987)]; N. G. Ptitsina *et al.*, Phys. Rev. **B56**, 10089 (1997).

[6] G. Bergman, Phys. Rev. **B3**, 3797 (1971); H. Takayama, Z. Phys. **263**, 329 (1973); S. G. Lisitsin, Sov. J. Low Temp. Phys. **1**, 728 (1975); S. J. Poon and T. H. Geballe, Phys. Rev. **233**, (1978).

[7] J. J. Lin and C. Y. Wu, Europhys. Lett. **29**, 141 (1995).

[8] A. Mittal, R. G. Wheeler, M. W. Keller, D. E. Prober, and R. N. Sacks, Surf. Sci. **361–362**, 537 (1996).

[9] G. Bergmann, Phys. Rep. **107**, 1 (1984).

Reduction in relaxation time due to ionized impurities in GaAs/AlGaAs quantum well structures

J. H. PARK[†], S. SENZAKI, N. MORI, C. HAMAGUCHI

Department of Electronic Engineering, Osaka University, 2-1 Yamada-oka, Suita, Osaka 565-0871, Japan

Time-resolved photoluminescence measurements in δ-doped GaAs/AlGaAs on the quantum well structures are performed to study effects of ionized impurities relaxation process of photoexcited carriers. It is theoretically shown that a thin quantum well with a δ-doping layer inserted in the barrier layer of double quantum wells enhances the impurity scattering rate significantly. Photoluminescence decay time in the δ-doped samples is found to decrease compared with the undoped samples.

© 2000 Academic Press

Key words: semiconductor quantum wells, electron tunneling, impurity scattering.

1. Introduction

A large nonlinear optical property is required for optical devices applications of quantum well (QW) structures [1]. Nonlinear optical properties of excitons in QW structures are especially remarkable because the absorption and refraction changes in the exciton peak due to the photo-excitation of carriers have a time constant shorter than 1 ps [2]. However, the slow decay time is a serious problem in these structures. From the viewpoint of optical device application of QW structures, it is desirable to achieve faster decay time of optically pumped carriers. To achieve a fast decay time utilizing a fast carrier tunneling time from one QW to another QW, longitudinal optical (LO) phonon and impurity scattering processes are widely and intensively investigated [3, 4].

In the present study, we propose a novel structure of double quantum wells (DQW) which enables us to control the transfer rate between the narrow well (NW) and the wide well (WW).

2. Theoretical background

We first theoretically demonstrate the fast impurity scattering time by making a thin QW in the barrier of DQW (TQW) and locating δ-doped impurities at the center of the thin QW. The impurity scattering rate for a nonresonant condition is proportional to a product of electron densities of initial and final states at the impurity position. In the limit of large detuning, the scattering rate can be written as [5],

$$\frac{\hbar}{2\tau_{\text{imp}}} = \frac{\pi^2}{2m^*} N_{\text{d}} \left(\frac{\hbar e^2}{2\pi\epsilon}\right)^2 \frac{1}{(E_n - E_{n'})^2} |\phi_{n'}^*(z_i)\phi_n(z_i)|^2, \tag{1}$$

[†]E-mail: park@ele.eng.osaka-u.ac.jp
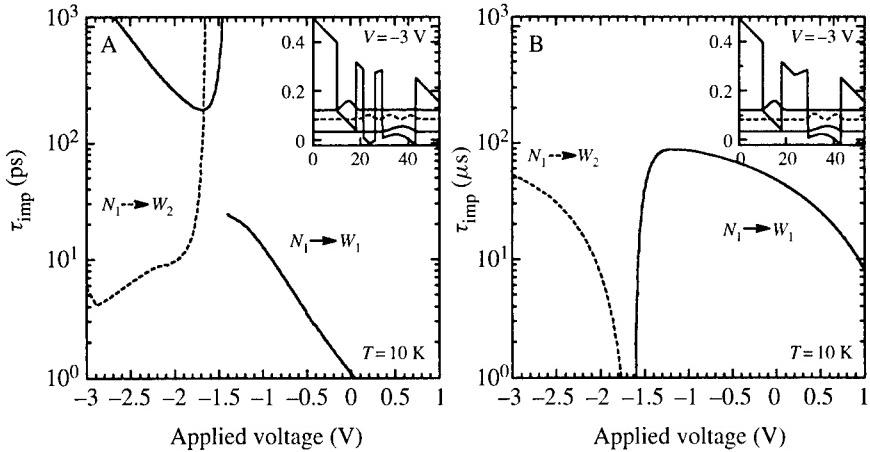
Fig. 1. Impurity scattering time, τ_{imp}, as a function of applied voltage, V, for A, δ-doped triple quantum wells and B, δ-doped double quantum wells. The areal donor density in $N_d = 2 \times 10^{12}$ cm^{-2}. Solid lines (dashed lines) represent τ_{imp} from the ground subband in the narrow well to the ground (first excited) subband in the wide well. The insets show a self-consistent electronic states at $V = -3$ V.

where N_d is the areal donor density, $E_n(E_{n'})$ is the initial (final) subband energy, $\phi_n(z)(\phi_{n'}(z))$ is the initial (final) state envelope function, and z_i is the impurity position.

Figure 1A shows calculated τ_{imp} as a function of externally applied voltage, V, for a δ-doped TQW. τ_{imp} is evaluated by using a self-consistent electronic state, which is shown in the inset of Fig. 1B for $V = -3$ V. In Fig. 1B we plot τ_{imp} for a δ-doped DQW comparison. As can be seen in Fig. 1, τ_{imp} for the structures with the thin QW (TQW) is considerably shorter than that for the structures without the thin QW (DQW).

3. Samples

We have grown the following four different types of GaAs/Al$_{0.35}$Ga$_{0.65}$As QW structures by molecular beam epitaxy (MBE) on GaAs substrates: (1) undoped DQW with a 14 nm thick WW, an 11 nm thick barrier layer and an 8 nm thick NW, (2) δ-doped DQW with a Si δ-doping layer at the center of the barrier, (3) undoped triple quantum wells (TQW) fabricated by replacing the 11 nm thick barrier of the undoped DQW with a QW structures consisting of a 3 nm thick Al$_{0.35}$Ga$_{0.65}$As barrier, a 5 nm thick thin QW of GaAs, and a 3 nm thick Al$_{0.35}$Ga$_{0.65}$As barrier, and (4) δ-doped TQW with a Si δ-doping layer at the center of the thin QW of the undoped TQW.

4. Results and discussion

We performed time-resolved PL experiments for the NW. Figure 2 shows the experimental results of PL decay time as a function of applied bias voltage from -3.5 to $+1.0$ V with 0.1 V step at 10 K. The decay time, τ_{decay}, for the δ-doped TQW is found to be faster than that for the other samples. A dip at $V \sim -1.5$ V in Fig. 2 can be seen for the TQW samples. It can be attributed to the fact that the first excited subband in WW coincide with the ground subband in NW. However, the decay time of DQW samples hardly change as a function of applied bias because of the thick barrier.

To confirm that the decrease in the decay time in the δ-doped TQW is due to the impurity scattering, we have grown three δ-doped TQW samples with different δ-doping densities. The temporal variation of the PL intensity for the NW of the three δ-doped TQW samples at 10 K is shown in Fig. 3. The decay time, τ_{decay}, is also found to become faster with increasing the δ-doping density.
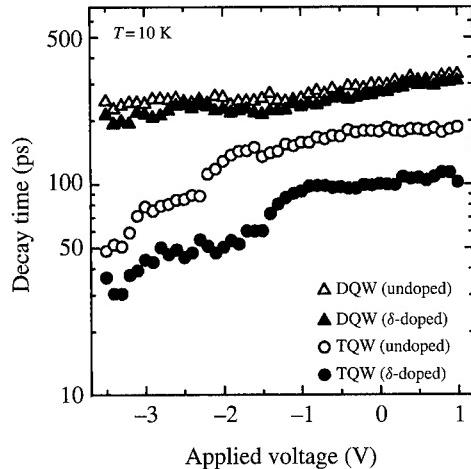
Fig. 2. Photoluminescence decay time from the narrow well in the undoped double quantum wells (DQW) (open triangles), the δ-doped DQW (closed triangles), the undoped triple quantum wells (TQW) (open circles) and the δ-doped TQW (closed circles) as a function of applied bias voltage from -3.5 to $+1.0$ V with a 0.1 V step at 10 K.
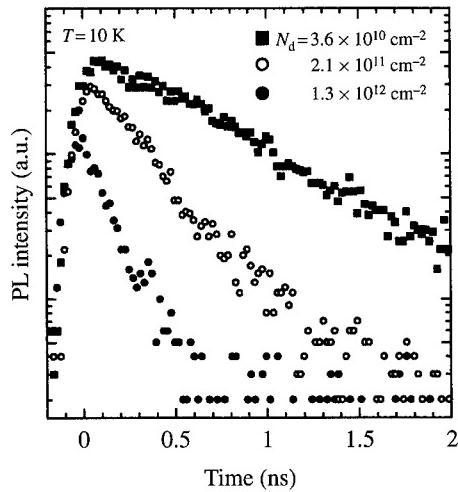
Fig. 3. Temporal variation of photoluminescence intensity for the narrow well of δ-doped triple quantum well samples with different δ-doping densities.

5. Summary

We proposed a novel structure of DQW with a thin QW with a δ-doping layer, which enables us to control the transfer rate between the NW and WW. Theoretical calculations show that a typical scattering time, τ_{imp}, is a few ps for TQW with δ-doping density of $N_d = 2 \times 10^{12}$ cm^{-2}. The PL decay time, τ_{decay}, in the δ-doped TQW is found to decrease by upto 45% compared with the undoped TQW. In δ-doped TQW samples, the decay time, τ_{decay}, becomes faster with increasing δ-doping density in TQW.

References

[1] A. Migus, A. Antonetti, D. Hulin, A. Mysyrowicz, H. M. Gibbs, N. Peyghambarian, and J. L. Jewell, Appl. Phys. Lett. **46**, 70 (1985).

[2] W. H. Knox, R. L. Fork, M. C. Downer, D. A. B. Miller, D. S. Chemla, C. V. Shank, A. C. Gossard, and W. Wiegmann, Phys. Rev. Lett. **54**, 1306 (1985).

[3] M. Stroscio, M. Kisin, G. Belenky, and S. Luryi, Appl. Phys. Lett. **75**, 3258 (1999) and references cited therein.

[4] J. H. Park, S. Ozaki, N. Mori, and C. Hamaguchi, Superlatt. Microstruct. **25**, 445 (1999) and references cited therein.

[5] R. Ferreira and G. Bastard, Rep. Prog. Phys. **60**, 345 (1997).

Chopped InGaAs/InP quantum wells for a polarization-independent space switch at 1.53 μm

B. H. P. Dorren, A. Yu. Silov, J. E. M. Haverkort, M. R. Leys

COBRA Inter-University Research Institute, Physics Department, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

D. H. P. Maat, Y. Zhu, F. H. Groen

Department of Applied Physics, Delft University of Technology, Delft, The Netherlands

J. H. Wolter

COBRA Inter-University Research Institute, Physics Department, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

We realized a polarization-independent Mach–Zehnder interferometric switch at 1.53 μm using chopped quantum wells. The calculated electroabsorption and electrorefraction agree well with experiment. The window for polarization-independent switching within 0.1 V is 12 nm. A phase shifting section of 1.7 mm yields a <1dB electroabsorption unbalance and a 6 V switching voltage.

© 2000 Academic Press

Key words: chopped quantum wells, interferometric space switcher.

Introduction

Future integrated optical cross connects for telecommunication applications require compact, low loss and polarization-independent Mach–Zehnder interferometric space (MZI) switches. It is the objective of this article to investigate whether chopped quantum wells (CQW) exploiting the quantum confined Stark effect (QCSE) can meet the necessary requirements.

In conventional InGaAs/InP quantum wells (QWs), waveguide transparency limits the QW-thickness to below 35 Å. Since the quantum confined Stark effect steeply increases [1] with well thickness, the QCSE in such a conventional well is very small. We present chopped quantum wells that consist of three strongly coupled 27 Å wide, 0.65% or 0.85% tensile-strained InGaAs wells separated by 15 Å barriers, as shown in Fig. 1. The CQW combines a bandgap of 1400 nm for waveguide transparency at 1.55 μm with an effective well width of 110 Å for a large QCSE. These chopped quantum wells show a QCSE red shift as large as 90 nm at 130 kV cm^{-1}, which is 30 times larger than in a 35 Å In$_{0.53}$Ga$_{0.47}$As QW. The CQW structure, including the interface layers [2, 3] which are introduced during growth, is given in Table 1.

As compared with a conventional QW, the heavy-hole envelope functions in a CQW behave differently. In the absence of an electric field, the lowest three heavy-hole energy levels are almost degenerate; the ground state and the second excited level are only 12 meV apart. In the presence of an electric field, each heavy-hole
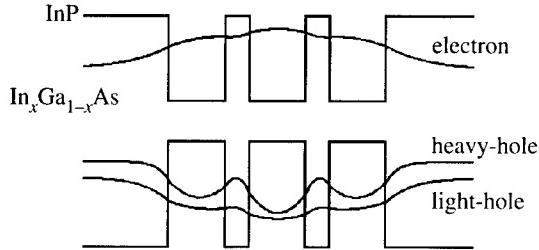
Fig. 1. Chopped quantum well with calculated electron and hole envelope functions (interface layers not shown).

Table 1: An overview of a single CQW with interfaces.

Layer	Thickness (Å)	Composition
Lower InP barrier	155	
InAs	3	
$In_xGa_{1-x}As$	24	$x = 0.56$ or 0.59
$In_xGa_{1-x}As_yP_{1-y}$	3	$x = 0.75$, $y = 0.5$
$InAs_yP_{1-y}$	12	$y = 0.12$
InAs	3	
$In_xGa_{1-x}As$	24	$x = 0.56$ or 0.59
$In_xGa_{1-x}As_yP_{1-y}$	3	$x = 0.75$, $y = 0.5$
$InAs_yP_{1-y}$	12	$y = 0.12$
InAs	3	
$In_xGa_{1-x}As$	24	$x = 0.56$ or 0.59
$In_xGa_{1-x}As_yP_{1-y}$	3	$x = 0.75$, $y = 0.5$
Upper InP barrier	155	

level is strongly localized in one of the InGaAs wells. The ground heavy-hole state localizes in the well that shows a large red shift, the first excited state localizes in the middle well and the second excited state localizes in the well which shows a blue shift. As a consequence, the net QCSE red shift in a CQW is two times larger than in a 110 Å $In_{0.6}Ga_{0.4}As_{0.85}P_{0.15}$ quantum well.

Calculated Mach–Zehnder switching features

To calculate the electroabsorption and electrorefraction in our CQW samples due to the QCSE, we use the 4 × 4 Luttinger–Kohn Hamiltonian. This method allows us to calculate the nonparabolic valence bands, the effective masses and optical matrix elements. Since the polarization properties of the QCSE strongly depend on the net strain in the sample, strained InAs and InGaAsP interfaces [4] are taken into account. We first calculated the absorption changes due to an applied electric field and subsequently calculated the electrorefraction spectra by taking a Kramers–Kronig transform.

Figure 2 shows the room-temperature absorption and electrorefraction spectra (inset) for the CQW sample at the built-in (13.5 kV cm^{-1}) and applied (46 kV cm^{-1}) field at the first switching point. The spectra were convoluted with a hyperbolic cosine function with a width of 45 meV, which was chosen equal to the photoluminescence linewidth. The calculated spectra agree well with our measured photocurrent spectra. We observe that the calculated electrorefraction for TM polarization is larger than for TE polarization.
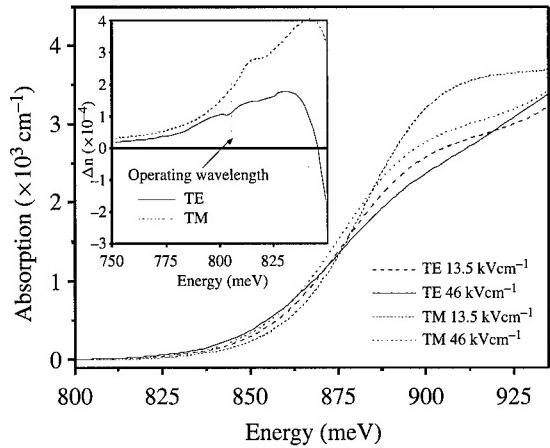
Fig. 2. Calculated TE and TM absorption spectra (300 K) for the 0.85% tensile-strained sample. The inset shows the calculated electrorefraction.
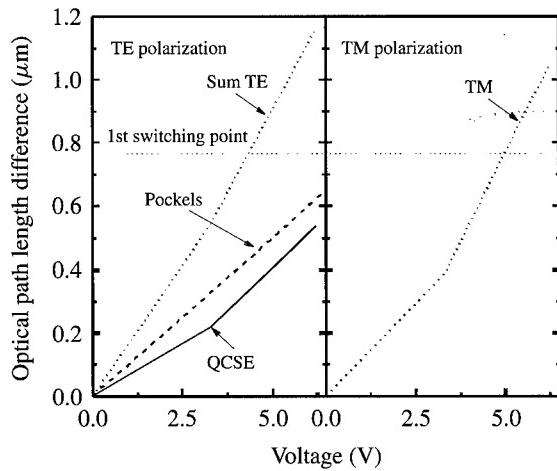

Fig. 3. Calculated optical path length difference due to the QCSE and the Pockels effect for TE and TM polarization for the 0.85% tensile-strained sample at 1530 nm.

Since the electrorefraction due to the QCSE is comparable to the Pockels effect, we have to add the electrorefraction due to the Pockels effect. The latter is positive for TE polarization for waveguides aligned along the $[1\bar{1}0]$ axis. The Pockels effect for TM polarization is zero. The electro-optic coefficient r_{41} for the intrinsic region of our structure is 1.5×10^{-12} m V^{-1}. The optical path length difference due to the Pockels effect in a switch with a 4 mm long phase shifting section is 0.1 μm V^{-1}. The optical power overlap Σ for the Pockels effect is unity since the waveguide mode propagates almost entirely in depleted material.

We are now able to calculate the optical path length difference $\rho n \cdot \Sigma \cdot l$ due to the QCSE and the Pockels effect using an optical power overlap of 0.42 between the waveguide mode and the quantum wells for the QCSE. The total optical path length as a function of voltage is shown in Fig. 3 for 1530 nm. These calculations indicate that it is possible to fabricate a polarization-independent switch with a switching voltage of 4.6 ± 0.3 V using CQWs. The optical path length difference of the Pockels effect and the QCSE are about equal at 1530 nm.
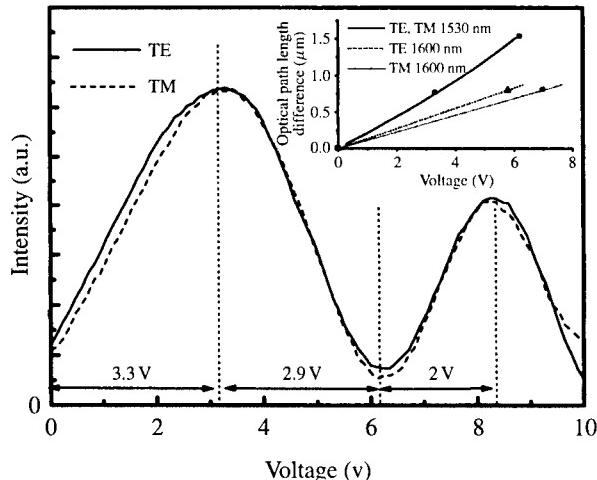
Fig. 4. Measured polarization-independent switching in a MZI using a 4 mm phase shifter with −0.85% strained chopped quantum wells (vertically scaled, see text).

Switching behavior and wavelength sensitivity of fabricated devices

Switches were fabricated using 0.65% and 0.85% tensile-strained CQW material. The switches consist of 2 × 2 multi-mode interference (MMI) couplers to divide the light equally over the 4 mm long phase shifting sections and to combine the light into one output waveguide. The MMI couplers are 713 μm long and 21 μm wide. The 3 μm wide waveguides are directed along the [1$\bar{1}$0] axis. Fabry Perot measurements at 1.55 μm show typical losses of 3.8 dB cm^{-1} for TE and 2 dB cm^{-1} for TM polarization.

Electroabsorption measurements were first performed in order to measure the added unbalance in the MZI due to electroabsorption in the phase section of the MZI. Unbalance in a MZI is known to be one of the sources for crosstalk; a 1 dB unbalance corresponds to −25 dB added crosstalk. For a 4 mm long phase shifting section with 0.85% tensile-strained CQWs, we observe that 1 dB electroabsorption occurs at 4.9 V for TE and 6.2 V for TM polarization. These values are in good agreement with our calculations. Polarization-independent switching is observed for the 0.85% strained sample at 1.53 μm as shown in Fig. 4 using 4 mm long phase sections. Owing to nonoptimal processing, the overall switching behavior is not polarization-independent. In this figure, the TE and TM switching curves are vertically scaled in order to account for the different losses. Clearly the switching voltages are polarization independent. The first switching point occurs at 3.3 ± 0.05 V for both polarizations.

The switching voltage for TM polarization is more sensitive to wavelength variations than the switching voltage for TE polarization since the electrorefraction due to the Pockels effect only contributes to the TE spectra. For the 0.65% tensile-strained sample, the switching voltage for TE and TM vary with wavelength with 0.042 V nm^{-1} and 0.050 V nm^{-1}, respectively (see Fig. 5). For the 0.85% tensile-strained sample, these numbers are 0.036 V nm^{-1} for TE and 0.053 V nm^{-1} for TM (see the inset in Fig. 4), which corresponds to polarization-independent switching voltages within ±0.1 V (−25 dB crosstalk penalty) in a 12 nm spectral window.

Minimum achievable dimensions and wavelength sensitivity

We have calculated the minimum achievable length of the phase shifting section, which adds 1 dB unbalance, or −25 dB crosstalk, due to electroabsorption at the first switching point. We plot the 1 dB elec-
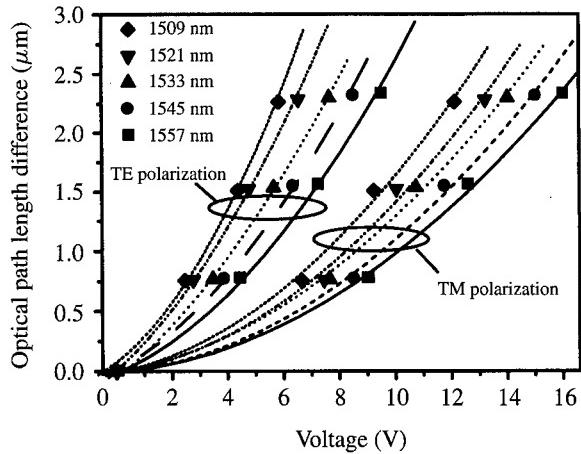
Fig. 5. Measured optical path length difference due to the QCSE for different operating wavelengths for the −0.65% strained CQW.
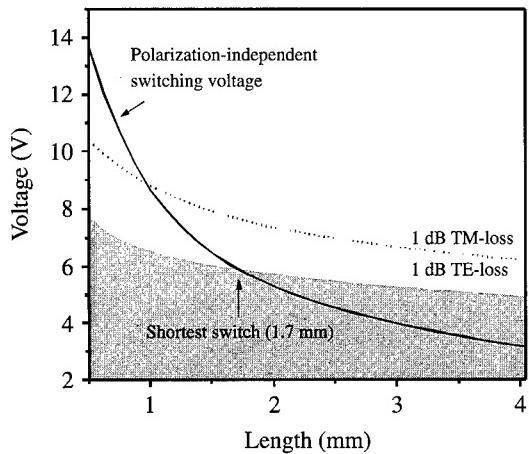

Fig. 6. Measured switching voltages together with the 1 dB loss curves versus the length of the phase section for the 0.85% tensile-strained sample at 1530 nm. The gray area indicates the allowable voltage–length combinations for less than 1 dB loss. A maximum of 1 dB unbalance requires a phase section longer than 1.7 mm.

troabsorption loss lengths versus voltage at 1530 nm in Fig. 6. The TE electroabsorption is the limiting loss mechanism. We also calculated the MZI switching voltages at 1530 nm as a function of the length of the phase shifting section. This curve is also shown in Fig. 6. The crossing of the 1 dB TE loss curve with the switching voltage line shows a minimum length of the phase shifting section of 1.7 mm.

The different wavelength dependence of the TE and TM switching voltages limits the spectral window for polarization-independent switching. We observe a shift of the switching voltage of 0.017 V nm^{-1} between both polarizations for the switch with 4 mm phase shifting sections. For a switch with a 1.7 mm phase shifting section, we expect a shift of the switching voltage of 0.022 V mm^{-1} between both polarizations. A polarization-independent switch with this length and a crosstalk penalty of −25 dB can be achieved over a spectral area as wide as 18 nm.

Conclusions

We realized an MZI at 1.53 μm employing InGaAs/InP chopped quantum wells with polarization-independent switching voltages. The calculated electroabsorption spectra have been checked by comparing them with photo-current measurements. It is essential to take the InAs and InGaAsP interface layers as well as the valence band interaction into account using the 4 $\times$ 4 Luttinger–Kohn Hamiltonian. The calculations show a polarization-independent switching voltage of 4.6 $\pm$ 0.3 V. The experimentally realized switching voltage is as low as 3.3 $\pm$ 0.05 V for a 4 mm phase shifter. For the 4 mm long switch the window for polarization-independent switching within $\pm$0.1 V, corresponding to a -25 dB crosstalk penalty, is 12 nm. Such chopped quantum wells can be applied for the fabrication of a polarization-independent switch with a phase shifter length of 1.7 mm, a switching voltage of 6 V and an 18 nm operating window.

References

[1] H. Temkin, D. Gershoni, and M. B. Panish, Appl. Phys.Lett. **50**, 1776 (1987).
[2] R. T. H. Rongen, A. J. C. van Rijswijk, M. R. Leys, C. M. van Es, H. Vonk, and J. H. Wolter, Semicond. Sci. Technol. **12**, 1 (1997).
[3] T. Marschner, J. Brübach, C. A. Verschuren, M. R. Leys, and J. H. Wolter, J. Appl. Phys. **83**, 3630 (1998).
[4] B. H. P. Dorren, A. Yu. Silov, M. R. Leys, J. E. M. Haverkort, and J. H. Wolter, in *Proceedings of the IEEE/LEOS Symposium Benelux Chapter, Gent, 1998*, edited by P. Demeester (University of Gent, Belgium, 1998) p. 133

Optical anisotropy of surface-emitting ZnSe/BeTe LEDs

G. Landwehr, D. R. Yakovlev[†], M. Keim, G. Reuscher, W. Ossau,
A. Waag

Physikalisches Institut der Universität Würzburg, 97074 Würzburg, Germany

A. V. Platonov, V. P. Kochereshko, E. L. Ivchenko

A.F. Ioffe Physical-Technical Institute, 194021 St Petersburg, Russia

The electric field-induced in-plane optical anisotropy (Pockels effect) of ZnSe/BeTe multi-layered heterostructures has been studied by analysing the linear polarization of the spatially indirect photoluminescence. A pronounced quantum-confined Pockels effect was found in ZnSe/BeTe double-barrier structures. A tight-binding model is proposed which consistently interprets the experimental findings and suggests a new mechanism for the effect as realized in type-II heterostructures. The model takes into account that the ZnSe/BeTe heterosystem exhibits a type-II band alignment with large band offsets and that in the zinc-blende lattice the chemical bonds are oriented in ⟨111⟩ directions and, when shifted along the [001] principal axis, alternatively change their orientation in the (001) plane from [1Ī0] to [110] and vice versa. A light-emitting diode based on a single ZnSe/BeTe interface is demonstrated.

© 2000 Academic Press

Key words: optical anisotropy, LED.

The in-plane optical anisotropy in quantum-well (QW) structures induced by normally applied electric fields is known as the quantum confined Pockels effect (QCPE). So far it has been observed in two III–V heterosystems with a type-I band alignment [1, 2]. Changes in polarization caused by an applied electric field are of the order of a few per cent. It has recently been shown that it is also possible to observe the Pockels effect in type-II semiconductor heterostructures [3]. Type-II structures have obvious advantages for optical studies of interface-related phenomena. In type-I QWs, the overlap between electron and hole wavefunctions extends over the whole QW. This averages and considerably reduces the effect of interface- induced and electric-field affected anisotropy. On the contrary, in type-II structures the radiative recombination is indirect in real space, it arises due to electron–hole overlap within an extremely narrow region adjacent to the interface. Consequently, the transition oscillator strength will be significantly affected by the orientation of the chemical bonds at the interface [4]. We report a very pronounced QCPE in ZnSe/BeTe double-barrier structures (DBS). The effect has been found experimentally and analysed theoretically. Modification of the band structure under applied external electric fields allows to detect signals from an individual interface and

---

[†]Also at: A.F. Ioffe Physical-Technical Institute, 194021 St Petersburg, Russia.
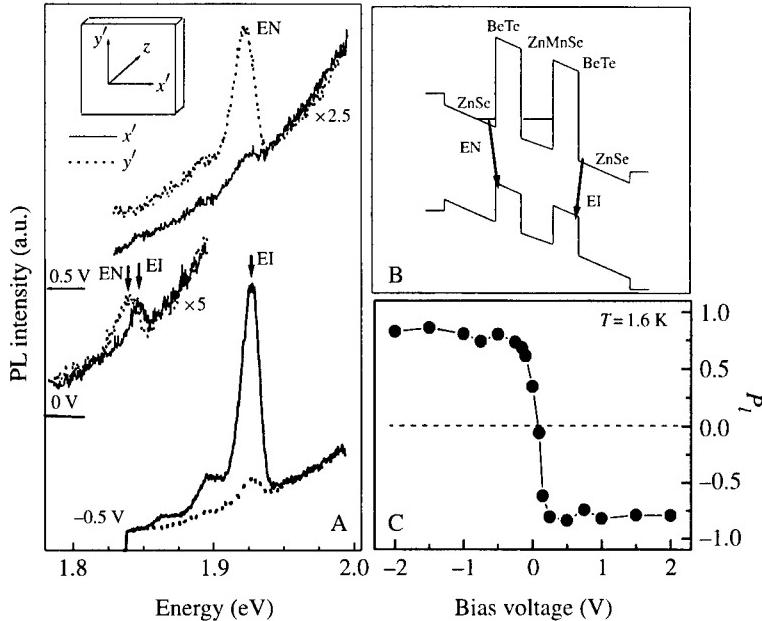
Fig. 1. A, The polarized PL spectra taken from a ZnSe/BeTe double-barrier structure at different bias voltages for the analyser orientation along the x' and y' axes ($T = 1.6$ K). B, The band diagram of the biased sample. C, Effect of an electric field on the polarization of the PL band arising due to optical transitions from the ZnSe emitter to the adjacent BeTe layer.

determine the orientation of chemical bonds from the field-induced optical anisotropy. We also found that the light which is emitted by the surface of ZnSe/BeTe light-emitting diodes is linearly polarized.

A set of resonant tunneling ZnSe/BeTe DBSs was grown on GaAs:Si(001) substrates by molecular-beam epitaxy. The layer sequence was as symmetric as possible and comprised of lattice matched BeZnSe:I (300 nm) contact layers, undoped ZnSe (25 nm) spacer layers, and the BeTe/$Zn_{0.9}Mn_{0.1}$Se/BeTe (4 nm/ 5 nm/4 nm) DBS (see Fig. 1B). The *normal* (BeTe-on-ZnSe) and *inverted* (ZnSe-on-BeTe) interfaces were grown, respectively, under Zn and Te termination and contained the Zn–Te chemical bonds. The PL spectra excited by UV lines of an Ar-ion laser and detected by a linear polarization along [1$\bar{1}$0] and [110] axes are shown in Fig. 1A. The field-induced optical anisotropy, i.e. Pockels effect, can be seen in Fig. 1C as the degree of linear polarization $P_l(I_{1\bar{1}0} - I_{110})/(I_{1\bar{1}0} + I_{110})$, where $I_{1\bar{1}0}$ and I_{110} are the integral intensities of the PL lines measured in the [1$\bar{1}$0] and [110] polarizer positions. Further details of the experimental realization of the QCPE in these type-II QWs have been published in Ref. [3].

The PL spectra detected in linear polarization along the [1$\bar{1}$0] and [110] axes are shown in Fig. 1A. The emission lines under study originate from the spatially indirect transitions involving electrons from ZnSe emitters and holes from BeTe layers (see Fig. 1B). At zero electric field a pair of PL lines at 1.84 eV arise from the recombination of electrons from the left- and right-hand side emitters and photoholes in BeTe (lines EN and EI). Positive electric fields push off electrons from the right Zn–Te interface and form a triangular potential well near the left Te–Zn interface. In this case the EN line increases in magnitude and the EI line almost vanishes. Reversal of the polarity of the applied voltage enhances the EI line intensity at the expense of the EN line. This allows us to select the luminescence emission of single interfaces and to switch the degree of polarization from +0.8 to −0.8 by changing the applied voltage from −0.5 to +0.5 V (see Fig. 1C).

In the ZnSe/BeTe system the conduction- and valence-band offsets amount to 2 and 1 eV, respectively, and the penetration depth of the electron wavefunction into the BeTe layer or for a hole into the ZnSe layer is about 3 Å. Therefore, in type-II direct-gap ZnSe/BeTe heterostructures the wavefunctions of an electron and
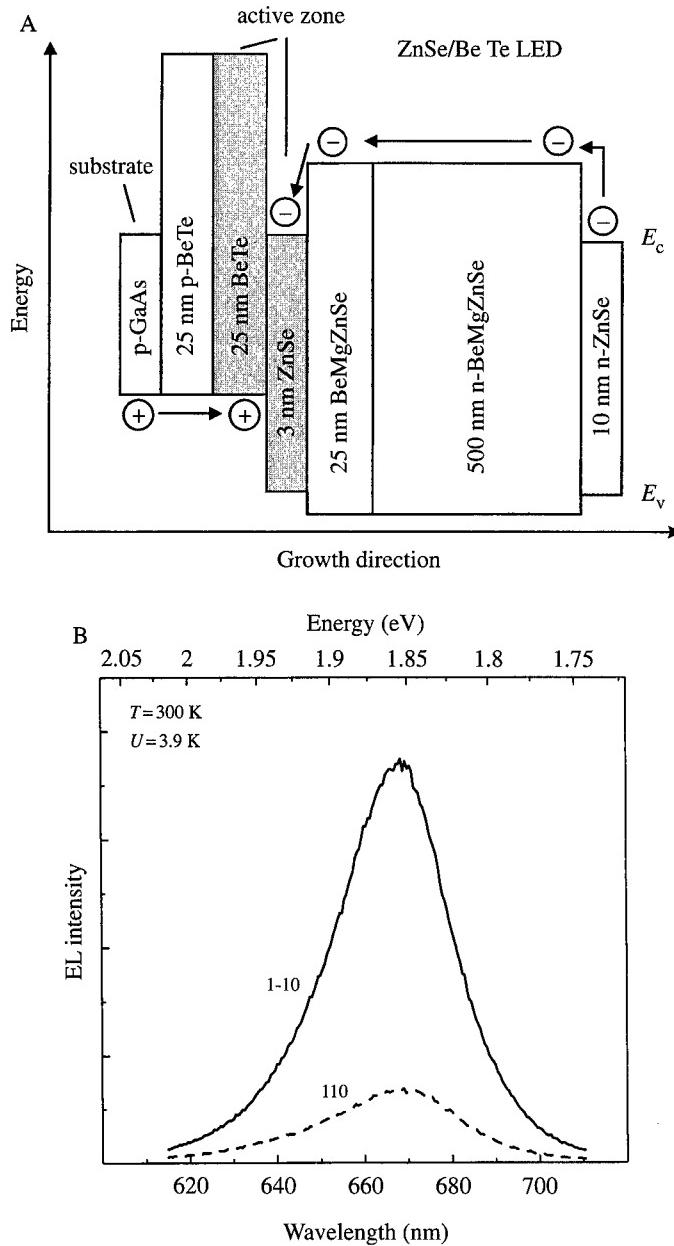
Fig. 2. A, Sketch of the band structure of a ZnSe/BeTe type-II LED. B, Surface-emitted electroluminescence signal detected in two linear polarizations along the [1$\bar{1}$0] and [110] crystallographic axes.

hole participating in the spatially indirect transition overlap substantially only over few atomic planes. In this case the calculation of the interband matrix elements requires the knowledge of the microscopic behavior of the wavefunctions at the interfaces which can be obtained by using the pseudo-potential or tight- binding approach. Application of the tight-binding model allows us to explain the experimental results and to relate the strong polarization degree to the orientation of the chemical bonds at the interface [3, 4].

The type-II heterosystem ZnSe/BeTe allows a new type of light-emitting diode (LED) which contains only a single ZnSe/BeTe interface. The emission of such a device can be adjusted between 640 and 515 nm [5]. It is based on the spatially indirect transition between the ZnSe conduction band and BeTe valence band. Normally one would not expect a device with a reasonable efficiency based on a type-II heterostructure. However, due to the strong carrier confinement at the interface at which the recombination of electrons and holes occurs efficiencies of 0.5% were achieved.

The devices were grown by molecular-beam epitaxy on p-type doped GaAs (100) substrates, which were provided with a thin GaAs buffer layer. In order to reduce the stacking fault density, four monolayers of BeTe were deposited before the growth of the diode. The growth was continued with 25 nm of 5×10^{18} cm^{-3} p-type BeTe. For the doping we used a nitrogen plasma at 300 W at partial pressure of 3×10^{-5} torr. Then the plasma was switched off and 25 nm undoped BeTe was grown followed by 3 nm of ZnSe acting as a quantum well for electrons. After a spacer of 25 nm undoped BeMgZnSe 500 nm lattice matched n-BeMgZnSe doped with ZnI$_2$ to 5×10^{17} cm^{-3} resulting in bandgaps ranging from 2.85 to 3.35 eV, respectively being grown. The structure was completed by 10 nm of n-ZnSe with an electron concentration of 1×10^{19} cm^{-3} for the ohmic metal contact. The scheme of the structure is presented in Fig. 2A.

Figure 2B shows electroluminescence emitted from the surface of a typical LED at room temperature. The emitted light is strongly (P_l is about 70%) linearly polarized in the direction [1$\bar{1}$0]. The linewidth (FWHM) of the spectra was about 100 meV which is about three times higher than in conventional ZnSe LEDs with ZnCdSe quantum wells. By narrowing the ZnSe quantum well the emission wavelength can be shifted towards shorter wavelengths, for a well thickness of 1 nm luminescence light of 515 nm is emitted. The total output was measured in an Ulbricht sphere. For a voltage of 4 V and a current of 15 mA the power was 0.3 mW, which corresponds to an external quantum efficiency of 0.5%. This high efficiency for a type-II transition can be explained by a carrier confinement at the type-II interface, which supports the validity of our theoretical approach. Degradation experiments did not show any decrease of efficiency after 1000 hours.

To conclude, in a resonant tunneling ZnSe/BeTe DBS with a type-II band alignment a pronounced quantum confined Pockels effect has been demonstrated. The theoretical analysis in the tight-binding model highlights the dominating role of the interface chemical bond orientation in the observed optical anisotropy. Therefore, the effect is a powerful tool for the investigation of the microscopic structure of heterointerfaces with monolayer resolution by means of nondestructive optical methods.

References

[1] S. H. Kwok, H. T. Grahn, K. Ploog, and R. Merlin, Phys. Rev. Lett. **69**, 973 (1992).

[2] O. Krebs, D. Rondi, J. L. Gentner, L. Goldstein, and P. Voisin, Phys. Rev. Lett. **80**, 5770 (1998).

[3] A. V. Platonov, V. P. Kochereshko, E. L. Ivchenko, G. V. Mikhailov, D. R. Yakovlev, M. Keim, W. Ossau, A. Waag, and G. Landwehr, Phys. Rev. Lett. **83**, 3546 (1999).

[4] D. R. Yakovlev, E. L. Ivchenko, V. P. Kochereshko, A. V. Platonov, S. V. Zeitsev, A. A. Maksimov, I. I. Tartakovskii, V. D. Kulakovskii, W. Ossain, M. Keim, A. Waag, and G. Landwehr, Phys. Rev. **B61**, R2421 (2000).

[5] G. Reuscher, M. Keim, H. J. Lugauer, A. Waag, and G. Landwehr, J. Cryst. Growth, in press.

sp3s* Tight-binding parameters for transport simulations in compound semiconductors

GERHARD KLIMECK[†], R. CHRIS BOWEN

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, U.S.A.

TIMOTHY B. BOYKIN

University of Alabama in Huntsville, Huntsville, AL 35899, U.S.A.

THOMAS A. CWIK

ECE Department and LICOS, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, U.S.A.

A genetic algorithm approach is used to fit orbital interaction energies of sp3s* tight-binding models for the nine binary compound semiconductors consistent of Ga, Al, In and As, P, Sb at room temperature. The new parameters are optimized to reproduce the bandstructure relevant to carrier transport in the lowest conduction band and the highest three valence bands. The accuracy of the other bands is sacrificed for the better reproduction of the effective masses in the bands of interest. Relevant band edges are reproduced to within a few meV and the effective masses deviate from the experimental values typically by less than 10%.

© 2000 Academic Press

Key words: sp3s* tight-binding, genetic algorithm, effective mass, parametrization, III–V material, GaAs, InAs, AlAs, GaP, InP, AlP, GaSb, InSb, AlSb.

1. Introduction

Nano-scaled electronic devices are characterized by material and charge density variations on the length scale of a few atoms. Tight-binding models [1] can resolve spatial material variations on an atomic scale and they bear the full crystalline and electronic symmetry of semiconductor materials in them. This ability has led to an increased use of these tight-binding models for the simulation of nano-scaled electronic devices (see references in [2]). While the tight-binding approach is systematically appealing, it bears a big problem in that the basic building constructs for the tight-binding Hamiltonian are not conduction band edges and effective masses, but orbital interaction energies (15 'free' parameters in the sp3s* model). These interaction energies are related to the global band structure and effective masses in a nontrivial manner [3].

It is important to realize that the tight-binding models do not include all the physics of electronic structure. The accuracy of these models strongly depends on the choice of orbitals that are included and the parametrization of the orbital interaction energies. It is for example understood, but not widely appreciated,

---

[†]E-mail: gekco@jpl.nasa.gov. Web: http://hpc.jpl.nasa.gov/PEP/gecko

Table 1: Bandstructure model parameters. All energies are in units of eV and the lattice constant is in units of nm.

Parameter	GaAs	AlAs	InAs	GaP	AlP	InP	GaSb	AlSb	InSb
lattice/(nm)	0.56660	0.56600	0.60583	0.54509	0.54635	0.58687	0.60959	0.61355	0.60583
$E(s,a)$	−3.53284	−3.21537	−9.57566	−8.63163	−8.93519	−7.91404	−7.16208	−4.55720	−7.80905
$E(p,a)$	0.27772	−0.09711	0.02402	0.77214	1.13009	0.08442	−0.17071	0.01635	−0.14734
$E(s,c)$	−8.11499	−9.52462	−2.21525	−1.77800	0.06175	−2.76662	−4.77036	−4.11800	−2.83599
$E(p,c)$	4.57341	4.97139	4.64241	4.17259	4.55816	4.75968	4.06643	4.87411	3.91522
$E(s^*,a)$	12.33930	12.05550	7.44461	11.90050	12.82470	9.88869	7.32190	9.84286	7.43195
$E(s^*,c)$	4.31241	3.99445	4.12648	7.99670	9.41477	7.66966	3.12330	7.43245	3.54540
$V(s,s)$	−6.87653	−8.84261	−5.06858	−7.21087	−6.68397	−6.16976	−6.60955	−6.63365	−4.89637
$V(x,x)$	1.33572	−0.01434	0.84908	1.83129	2.28630	0.75617	0.58073	1.10706	0.75260
$V(x,y)$	5.07596	4.25949	4.68538	4.87432	5.12891	4.23370	4.76520	4.89960	4.48030
$V(sa,pc)$	2.85929	2.42476	2.51793	6.12826	9.44286	3.62283	3.00325	4.58724	3.33714
$V(sc,pa)$	11.09774	13.20317	6.18038	6.10944	5.93164	6.90390	7.78033	8.53398	5.60426
$V(s^*a,pc)$	6.31619	5.83246	3.79662	6.69771	10.08057	4.61375	4.69778	7.38446	4.59953
$V(pa,s^*c)$	5.02335	4.60075	2.45537	6.33303	4.80831	6.18932	4.09285	6.29608	−2.53756
Δ_a	0.32703	0.29145	0.38159	0.05379	0.04600	0.09400	0.75773	0.70373	0.85794
Δ_c	0.12000	0.03152	0.37518	0.21636	0.01608	0.54000	0.15778	0.03062	0.51000

that the sp3s* nearest-neighbor model pathologically predicts an infinite transverse mass at the X-point [2, 3]. With the limitations of the sp3s* model in mind it must be emphasized that early parametrizations [1] provided more global band structure fits for bands which can be probed by optical measurements. The complication in the fitting process of the 15 orbital interaction energies to measurable quantities has led to the use of the seminal Vogl [1] parameters in areas in which they were not intended. For example, quantitative electron transport simulations in high-performance resonant tunneling diodes [4] require the proper representation of relative band energies as well as band curvature of the conduction *and* valence bands, even for a purely electronic device (no hole transport).

2. Method and results

To enable an automated fitting of the orbital interaction energies to measureable quantities such as masses and band edges, a genetic algorithm-based procedure has been developed [2]. At the core of this algorithm is a single-valued fitness function comprised of weighted standard deviations that are to be minimized for a list of band structure properties. The details of this algorithm are documented in [2]. The strength of the algorithm is its ability to search for global minima in a very nonlinear search space. In the course of this work it was verified that derivative-based search algorithms will get stuck in local minima.

The target material properties were taken from [1, 5, 6]. Reported experimental material properties were preferred over theoretical values. The split-off band parameters Δ were restricted to be within a window of ±30% of the parameters listed by Vogl [1]. Note that the Vogl parametrization does not include these parameters and therefore does not model the split-off bands properly. The other orbital interaction energies were allowed to float free in energy. The interaction energies are normalized such that the valence band edge at Γ corresponds to 0 eV. The fitting for the nine binary compounds was performed in nine independent procedures. That implies that the on-site energies of one constituent vary from compound to compound. The new sp3s* parameters are listed in Table 1 to five decimal digit accuracy in units of eV.

The resulting band structure properties are compared with the target properties in Tables 2 and 3. If no

Table 2: Material properties (target and tight binding (TB) computed with parameters in Table 1) for GaAs, AlAs, InAs, GaP, and AlP grouped into properties at Γ, X, L and other spectrally accessible bands. Properties at Γ, and band edges at X and L received the highest weights. The other spectrally accessible bands are weighed only nominally except for Γ_{7c}, L_{7c} and X_{7c}.

Property	GaAs target	GaAs TB	AlAs target	AlAs TB	InAs target	InAs TB	GaP target	GaP TB	AlP target	AlP TB
Γ_{6c}	1.424	1.424	3.020	3.018	0.370	0.368	2.780	2.779	3.620	3.620
Δ_{so}	0.340	0.312	0.300	0.291	0.380	0.381	0.080	0.080	0.040	0.040
m_Γ^*	0.067	0.068	0.150	0.154	0.024	0.024	0.077	0.219		0.249
$m_{lh}^*[001]$	-0.087	-0.080	-0.163	-0.151	-0.027	-0.028	-0.162	-0.160	-0.211	-0.184
$m_{lh}^*[011]$	-0.080	-0.073	-0.140	-0.131	-0.026	-0.027	-0.138	-0.142		-0.164
$m_{lh}^*[111]$	-0.079	-0.072	-0.135	-0.127	-0.026	-0.027	-0.133	-0.138	-0.145	-0.159
$m_{hh}^*[001]$	-0.403	-0.389	-0.516	-0.520	-0.345	-0.364	-0.448	-0.494	-0.513	-0.552
$m_{hh}^*[011]$	-0.660	-0.663	-1.098	-1.100	-0.639	-0.657	-0.853	-0.809		-0.864
$m_{hh}^*[111]$	-0.813	-0.838	-1.570	-1.578	-0.876	-0.883	-1.119	-0.982	-1.372	-1.026
m_{so}^*	-0.150	-0.159	-0.240	-0.262		-0.098	-0.465	-0.248		-0.279
Δ_{min}	≈ 0.90	0.860	≈ 0.75	0.800		1.000	≈ 0.90	0.784	1.000	1.000
E_c^Δ	1.900	1.900	2.170	2.171	2.280	2.345	2.272	2.272	2.505	2.505
m_{Xl}^*	1.300	1.301	1.100	1.006		1.103	0.910	0.910		3.052
m_{Xt}^*	0.230	3.990	0.190	2.009		inf	0.254	2.265		inf
L_{6c}	1.708	1.708	2.352	2.351	1.500	1.460	2.637	2.635	3.570	3.540
m_{Ll}^*	1.900	1.775	1.900	2.609		1.852		24.323		>100
m_{Lt}^*	0.075	0.713	0.096	0.860		0.304		2.306		<0
Γ_{6v}	-13.100	-13.072	-11.950	-15.758	-12.300	-12.159	-12.300	-13.189	-11.820	-12.493
Γ_{7c}	4.530	4.312	4.540	3.994	4.390	4.126	4.870	4.845	5.090	5.687
Γ_{8c}	4.716	4.865	4.690	4.950	4.630	4.543	4.920	5.035	5.090	5.709
X_{5v}	-6.800	-4.710	-5.690	-4.064	-6.600	-7.595	-7.070	-7.876	-7.000	-6.085
X_{6v}	-2.880	-3.151	-2.410	-2.591	-2.400	-2.948	-2.730	-2.703	-2.270	-2.572
X_{7v}	-2.800	-3.023	-2.410	-2.449	-2.400	-2.836	-2.730	-2.678	-2.270	-2.555
X_{6c}	1.980	1.932	2.229	2.285	2.300	2.345	2.350	2.432	2.505	2.505
X_{7c}	2.320	2.117	3.800	2.461	2.500	2.849	2.750	2.651	4.300	4.461
L_{5v}	-8.000	-5.168	-6.000	-4.606	-6.230	-7.062	-6.840	-7.452	-6.000	-6.222
L_{6v}	-1.420	-1.528	-0.880	-0.956	-1.200	-1.399	-1.100	-1.320	-1.000	-1.253
L_{7v}	-1.200	-1.340		-0.782	-0.900	-1.144	-1.100	-1.254	-1.000	-1.228
L_{7c}	5.470	3.246	5.860	3.362	5.400	3.713	5.740	3.868	6.000	4.738

target value is listed, no 'trusted' experimental or theoretical values could be found in [1, 5, 6]. Only a few remarks relating to and beyond the results posted in Tables 2 and 3 are given due to space limitations. The light hole mass is more easily fitted in small gap materials, where the magnitude of the conduction band to split-off hole band coupling is greater than that of the light hole to higher conduction bands [3]. Conduction band minima and longitudinal masses near the X-point on the Δ line are fitted for several materials. Note that the sp3s* model breaks down as noted and visualized in [2, 3] and no attempts have been made to fit the transverse electron mass at that minimum point. Only nominal weights are associated with the longitudinal L-point mass. No weight is given to the transverse L-point mass. During the fitting process it was observed that the upper conduction bands will be pushed unphysically close to the first conduction band if only effective

Table 3: Same as Table 2 for InP, GaSb, AlSb, and InSb.

Property	InP target	InP TB	GaSb target	GaSb TB	AlSb target	AlSb TB	InSb target	InSb TB
Γ_{6c}	1.344	1.345	0.750	0.751	2.300	2.300	0.169	0.169
Δ_{so}	0.108	0.105	0.760	0.747	0.673	0.675	0.858	0.847
m^*_Γ	0.077	0.078	0.041	0.042	0.120	0.121	0.014	0.014
$m^*_{lh}[001]$	−0.096	−0.082	−0.050	−0.043	−0.123	−0.099	−0.015	−0.014
$m^*_{lh}[011]$	−0.087	−0.076	−0.046	−0.040	−0.100	−0.089	−0.015	−0.014
$m^*_{lh}[111]$	−0.085	−0.075	−0.045	−0.040	−0.091	−0.086	−0.015	−0.014
$m^*_{hh}[001]$	−0.474	−0.480	−0.285	−0.300	−0.336	−0.363	−0.278	−0.287
$m^*_{hh}[011]$	−0.933	−0.886	−0.551	−0.559	−0.500	−0.632	−0.525	−0.531
$m^*_{hh}[111]$	−1.300	−1.187	−0.763	−0.759	−0.872	−0.800	−0.737	−0.732
m^*_{so}	−0.120	−0.150		−0.134	−0.290	−0.196		−0.132
Δ_{min}	≈0.90	0.900	≈0.80	0.850	≈0.80	0.849		1.000
E_c^Δ	2.304	2.312	1.180	1.181	1.615	1.615	1.493	1.524
m^*_{Xl}		5.985	1.510	1.424	1.800	1.576		1.181
m^*_{Xt}		11.452	0.220	3.183	0.260	2.734		inf
L_{6c}	1.954	1.958	0.832	0.833	2.211	2.211	0.931	0.930
m^*_{Ll}		2.588	0.950	1.421		24.866		1.838
m^*_{Lt}		0.636	0.110	0.405		1.125	0.090	0.312
Γ_{6v}	−11.000	−12.025	−12.000	−12.683	−11.100	−10.975	−11.730	−10.814
Γ_{7c}	4.640	4.527	3.400	3.123	3.740	5.075	3.370	3.545
Γ_{8c}	4.920	5.055	4.700	4.033	4.000	5.135	3.740	3.703
X_{5v}	−6.010	−8.957	−6.760	−8.010	−6.760	−6.995	−6.430	−7.498
X_{6v}	−2.090	−2.440	−3.000	−3.433	−3.000	−3.192	−2.450	−3.190
X_{7v}	−2.060	−2.392	−2.500	−3.107	−2.500	−2.858	−2.240	−2.890
X_{6c}	2.304	2.314	1.400	1.210	1.615	1.632	1.493	1.524
X_{7c}	2.970	2.913		1.396	3.020	2.614	1.830	2.038
L_{5v}	−5.840	−7.774	−6.250	−7.689	−6.250	−7.160	−5.920	−7.054
L_{6v}	−1.090	−1.054	−1.450	−1.682	−1.450	−1.613	−1.400	−1.698
L_{7v}		−0.942	−1.000	−1.249	−1.000	−1.225	−0.900	−1.165
L_{7c}	5.580	4.232	4.400	2.432	4.400	3.476	4.000	2.969

masses and band edges of the lowest conduction and three highest valence bands are considered. Including the Γ_{7c}, L_{7c} and X_{7c} with higher than nominal weights keeps the upper conduction bands well above the lowest conduction band. The agreement with the L_{6c} value must be driven towards cautiously as a minima on the Λ line starts to develop instead of the minimum at the L-point.

3. Summary

This work provides a new parametrization of the widely used sp3s* tight-binding model for the nine binary compounds consisting of In, Ga, Al and Sb, As, P. The new parametrization is expected to be more suitable for nano-electronic transport simulations than the seminal work by Vogl *et al.* [1] since it focuses on the accurate modeling of effective masses and central band-edges. This is, to our knowledge, the first attempt to systematically fit anisotropic hole masses at Γ as well as electron masess at Γ and X in the sp3s* model.
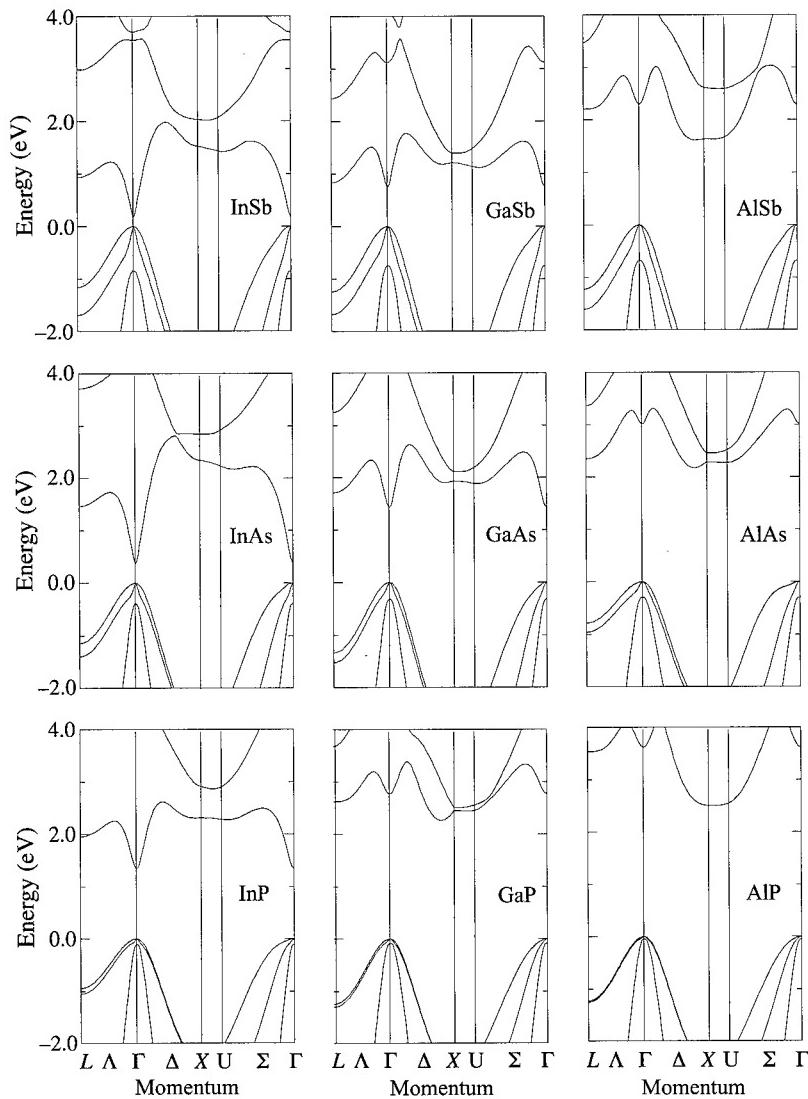
Fig. 1. Dispersion relations for nine binary semiconductor compounds computed with the parameters listed in Table 1.

References

[1] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids **44**, 365 (1983).

[2] G. Klimeck *et al.*, Superlatt. Microstruct. **27**, 77 (2000).

[3] T. Boykin, G. Klimeck, R. C. Bowen, and R. Lake, Phys. Rev. **B56**, 4102 (1997).

[4] R. C. Bowen *et al.*, J. Appl. Phys. **81**, 3207 (1997).

[5] O. Madelung, *Semiconductors—Basic Data* (Springer, Berlin, 1996).

[6] J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani, Phys. Rev. **B57**, 6493 (1998).

Cyclotron masses in InGaAs/GaAs superlattices and InGaAs/AlAs superlattices

H. MOMOSE[†], S. UEHARA, N. MORI, C. HAMAGUCHI

Department of Electronic Engineering, Osaka University, Suita, Osaka 565-0871, Japan

H. ARIMOTO, T. IKAIDA, N. MIURA

Institute for Solid State Physics, University of Tokyo, Roppongi, Tokyo 106-8666, Japan

Cyclotron resonance (CR) measurements have been carried out to evaluate the effective mass of electron in $(InGaAs)_n/(GaAs)_n$ superlattices (SLs) and $(InGaAs)_n/(AlAs)_n$ SLs. To clarify the dependence of cyclotron mass on the monolayer number n, we measured CR signals using pulsed high-magnetic fields up to 150 T and a far-infrared laser. We found clear cyclotron resonances in the transmission of 10.6 μm at 75 T at room temperature in $(InGaAs)_n/(GaAs)_n$ SLs and little dependence on the monolayer number n in the SLs. However, for $(InGaAs)_n/(AlAs)_n$ SLs, a large dependence of cyclotron mass on the monolayer number n was observed. We consider that these dependencies are related to the difference between the barrier height in the SLs and the influence of nonparabolicity on the conduction subbands in the SLs.

© 2000 Academic Press

Key words: cyclotron mass, InGaAs/GaAs superlattice, InGaAs/AlAs superlattice.

1. Introduction

Recently, progress in pulsed high-magnetic field technology has enabled us to perform accurate measurements of the properties of matter in very high-magnetic fields [1]. Cyclotron resonance (CR) is one of the basic tools for studying the electronic properties of physical systems in an external magnetic field. Since the high field makes the cyclotron energy $\hbar\omega_c$ so high and the cyclotron orbital $(\hbar/eB)^{1/2}$ so small, megagauss fields allow us to study various properties of semiconductor nanostructures with very high resolution and with a significant interaction between the magnetic field and the artificially introduced quantum potentials. Thus, we have studied InGaAs/GaAs superlattices (SLs) and InGaAs/AlAs SLs using CR measurements under high-magnetic fields. Properties of InGaAs/GaAs system and InGaAs/AlAs system have been investigated by many researchers, and many high-speed electronic and opt-electronic devices have been designed and developed [2, 3]. However, many fundamental properties, such as electron effective mass, band offset, and so on, are not clearly understood. In this paper, we report CR measurements in InGaAs/GaAs SLs under pulsed high-magnetic fields up to 150 T and evaluate the cyclotron mass in the SLs as a function of the monolayer number n.

[†]Author to whom correspondence should be addressed. E-mail:momose@ele.eng.osaka-u.ac.jp
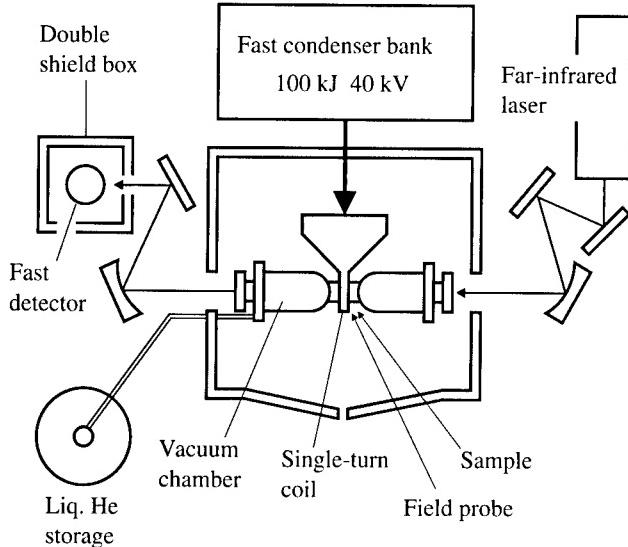
Fig. 1. Block diagram for CR measurement. A pulsed high-magnetic field up to 150 T is produced by using a single-turn coil technique. The magnetic field is applied parallel to the growth direction of SLs.

2. Experimental procedures

The samples of $(In_{0.2}Ga_{0.8}As)_n/(GaAs)_n$ SLs and $(In_{0.2}Ga_{0.8}As)_n/(AlAs)_n$ SLs were grown by molecular beam epitaxy on (001)-oriented semi-insulating GaAs substrates. To supply electrons in the SL region, the samples were delta-doped with Si in the center of each InGaAs well.

The applied magnetic field up to 150 T was generated by using single-turn coil technique [4]. A block-diagram of the experimental apparatus is shows in Fig. 1. The magnetic fields were produced by supplying a large current ($\sim$2.5 MA) to thin single-turn coil from a fast condenser bank with a charged energy of 100 kJ at 40 kV. The generated magnetic fields were measured by a pick-up coil wound on the sample. The duration time of the magnetic pulse was about 7 μs. The CR signals were observed in transmission from a far-infrared pulse laser synchronized with the magnetic field. In this study, we used a CO_2 laser with a wavelength of 10.6 μm as a far-infrared laser and the measurements were done at room temperature to observed free-electron CR signals avoiding impurity CR signals [5]. To estimate the electron mass in the planes of the SLs, we applied the magnetic field parallel to the growth direction of the SLs.

3. Experimental results and discussion

Figure 2 shows the CR signals of $(InGaAs)_n/(GaAs)_n$ SLs with $n = 6$–18 in the transmission of 10.6 μm radiation at room temperature. We obtained the cyclotron mass from the resonance position by fitting a CR curve with a Lorentian curve. We can find resonance peaks for $(InGaAs)_n/(GaAs)_n$ SLs in CR signals at 75 T, which gives the cyclotron mass of $0.074m_0$, and little dependence of the cyclotron mass on the monolayer number n. This dependency of the cyclotron mass is in contrast to the change of the effective mass in GaAs/AlAs multi-quantum wells [6].

It is well known that the effective mass of bulk GaAs is about $0.067m_0$ and the effective mass of bulk $In_{0.2}Ga_{0.8}As$ is lighter than bulk GaAs ($m^*_{In_{0.2}Ga_{0.8}As} \sim 0.059m_0$). In our experiment, the cyclotron masses of $(InGaAs)_n/(GaAs)_n$ SLs with $n = 6 \sim 18$ were $0.074m_0$, which is heavier than the effective mass of bulk material in the well and barrier. We consider that the these results are due to the influence of nonparabolicity on the conduction subbands in the SLs [7].
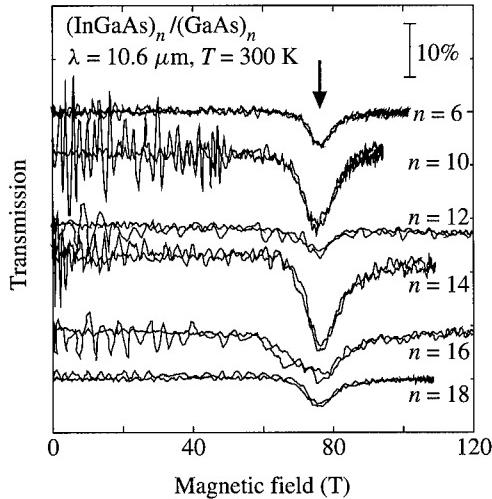
Fig. 2. Experimental results of CR for $(InGaAs)_n/(GaAs)_n$ SLs with $n = 6$–18 in the transmission of 10.6 μm radiation at room temperature. The resonance at around 75 T gives the cyclotron mass of $0.074m_0$.
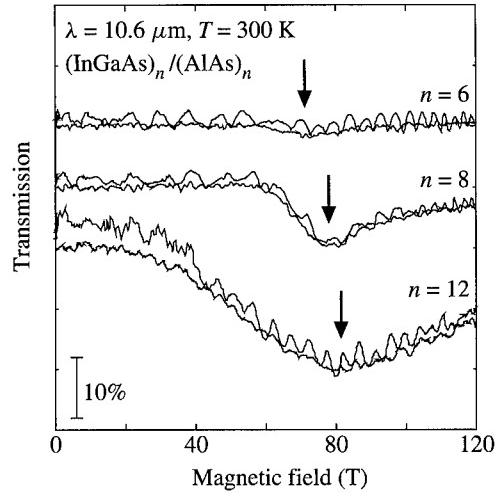
Fig. 3. Experimental results of CR for $(InGaAs)_n/(AlAs)_n$ SLs with $n = 6$–12 in the transmission of 10.6 μm radiation at room temperature. The cyclotron masses of the SLs with $n = 6, 8, 12$ were $0.073m_0$, $0.078m_0$, $0.080m_0$, respectively.

Figure 3 shows the CR signals of $(InGaAs)_n/(AlAs)_n$ SLs with $n = 6$–12 in the transmission of 10.6 μm radiation at room temperature. We can find resonance peaks at 74 T for $n = 6$, 78 T for $n = 8$ and 81 T for $n = 12$. From these results, the cyclotron masses of $(InGaAs)_n/(AlAs)_n$ SLs with $n = 6, 8, 12$ were estimated at $0.073m_0$, $0.078m_0$, $0.080m_0$, respectively. This dependency of the cyclotron mass in $(InGaAs)_n/(AlAs)_n$ SLs is in contrast to the near independency of $(InGaAs)_n/(GaAs)_n$ SLs.

Estimated cyclotron masses in $(InGaAs)_n/(GaAs)_n$ SLs ($\bullet$) and $(InGaAs)_n/(AlAs)_n$ SLs ($\circ$) are plotted in Fig. 4. The difference in behavior of cyclotron mass between $(InGaAs)_n/(GaAs)_n$ SLs and $(InGaAs)_n/(AlAs)_n$ SLs is caused by quantization of electrons in the SLs. Because of the extremely high barrier in the $(InGaAs)_n/(AlAs)_n$ SLs compared with the $(InGaAs)_n/(GaAs)_n$ SLs, the influence of the quantization of electrons in the SLs is reflected in the effective mass. According to the calculation of Ekenberg [8],
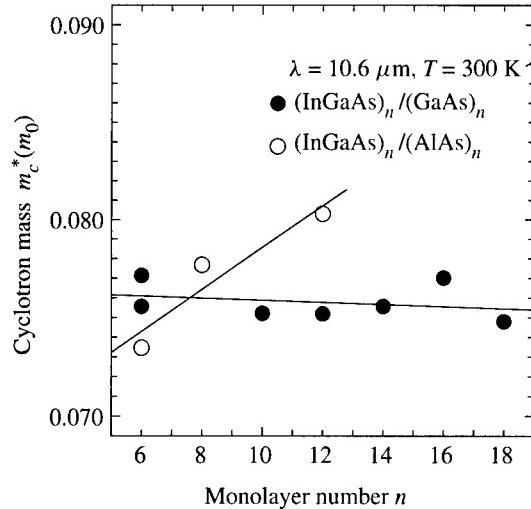
Fig. 4. Estimated cyclotron masses in $(InGaAs)_n/(GaAs)_n$ SLs (●) and $(InGaAs)_n/(AlAs)_n$ SLs (o). A pulsed high-magnetic field up to 150 T is produced by using a single-turn coil technique. The magnetic field is applied parallel to the growth direction of SLs.

the cyclotron mass becomes heavier with decreasing well width in quantum well structure. However, our results of the cyclotron mass in $(InGaAs)_n/(AlAs)_n$ SLs is in conflict with Ekenberg's theoretical prediction. Although we neglect strain effects in SLs, for more detailed analysis the strain effects become important factor in the behavior of cyclotron masses.

In summary, we have studied the cyclotron mass in $(InGaAs)_n/(GaAs)_n$ SLs and $(InGaAs)_n/(AlAs)_n$ SLs. Clear cyclotron resonance was observed and the cyclotron mass was estimated $0.074m_0$ in $(InGaAs)_n/(GaAs)_n$ SLs. We also found little dependence on the monolayer number n in the SLs. However, for $(InGaAs)_n/(AlAs)_n$ SLs, a large dependence of cyclotron mass on the monolayer number n was obtained. We consider that these dependencies are due to a difference of barrier height and effects of nonparabolic conduction subbands in the SLs.

References

[1] N. Miura, Physica **B201**, 40 (1994).

[2] H. Morkoc and H. Unlu, *Semiconductors and Semimetals*, edited by R. Dingle (Academic, New York, 1987) p. 135

[3] R. M. Kapre, W. T. Tsang, M. C. Wu, and Y. K. Chen, Proc. SPIE **1788**, 37 (1993).

[4] K. Nakao, F. Herlach, T. Goto, S. Takeyama, T. Sakakibara, and N. Miura, J. Phys. E: Sci. Instrum., **18**, 1018 (1985).

[5] H. Momose, N. Mori, C. Hamaguchi, T. Ikaida, H. Arimoto, and N. Miura, Mater. Sci. Eng. **B64**, 137 (1999).

[6] J. G. Michels, R. J. Warburton, R. J. Nicholas, J. J. Harris, and C. T. Foxon, Physica **B184**, 159 (1993).

[7] N. Kotera, Y. Shimamoto, H. Arimoto, N. Miura, K. Tanaka, and T. Mishima, Superlatt. Microstruct. **24**, 255 (1998).

[8] U. Ekenberg, Phys. Rev. **B40**, 7714 (1989).

Quantum mechanical Monte Carlo approach to electron transport at heterointerface

HIDEAKI TSUCHIYA, TANROKU MIYOSHI

Department of Electrical and Electronics Engineering, Kobe University, 1-1, Rokko-dai, Nada-ku, Kobe 657-8501, Japan

With the progress of LSI technology, the electronic device size is scaled down to the sub 0.1 μm region. In such an ultrasmall device, it is indispensable to take quantum mechanical effects into account in device modeling. In this paper, we present a newly developed quantum Monte Carlo device simulation applicable to ultrasmall semiconductor devices. In this model, the quantum effects are represented in terms of quantum mechanically corrected potential in the classical Boltzmann equation. It is demonstrated that the quantum transport effects such as tunneling and energy quantization in ultrasmall semiconductor devices are obtained for the first time by using the standard Monte Carlo techniques.

© 2000 Academic Press

Key words: tunneling, energy quantization, Monte Carlo simulation, quantum correction of potential.

1. Introduction

Semiconductor device integration, typically in LSI circuits, has progressed rapidly in recent years. With the downsizing of the MOSFETs, the hot carrier and quantum mechanical effects have become more and more important. The Monte Carlo method, which is a general statistical numerical method for solving the Boltzmann equation directly, is a powerful tool for investigating electron transport in semiconductor devices. To date, the Monte Carlo simulation is considered to be a well-established method in which the full band structure of crystals, carrier Coulomb interactions, details of scattering mechanisms and evaluations of impact ionization thresholds based upon the realistic band structure have been successfully implemented [1]. However, the Boltzmann equation as a basis for all Monte Carlo simulations is a fully classical equation though collisions are assumed as localized events in space and time and calculated quantum mechanically. Thus, the conventional Monte Carlo techniques have been unable to embody the carrier quantum transport effects such as tunneling and energy quantization.

Recently, we have formulated a novel Boltzmann-like quantum transport equation based upon the Wigner distribution function [2]. In this model, the quantum effects are represented in terms of quantum mechanically corrected potential in the classical Boltzmann equation. In this paper, the Monte Carlo method is used as a direct solver of the quantum transport equation. We will demonstrate that the quantum transport effects of ultrasmall semiconductor devices are obtained for the first time by using the standard Monte Carlo techniques.

2. Quantum Monte Carlo method

The quantum transport equation for the Wigner distribution function, which corresponds to a quantum mechanical distribution function, is given in the form of a modified Boltzmann equation as [3]

$$\frac{\partial f}{\partial t} + \frac{\hbar k}{m^*} \cdot \nabla_r f - \frac{1}{\hbar} \nabla_r U \cdot \nabla_k f - \sum_{\alpha=1}^{\infty} \frac{(1/2i)^{2\alpha}}{\hbar(2\alpha+1)!} \nabla_r^{2\alpha+1} U \cdot \nabla_k^{2\alpha+1} f = \left(\frac{\partial f}{\partial t}\right)_C, \qquad (1)$$

where the fourth term on the left-hand side denotes the quantum correction due to the spatially varying potential energy $U(r)$, which accounts for various quantum effects. Incidentally, the quantum correction terms become zero in the classical limit of $\hbar \to 0$. Although the quantum transport eqn (1) provides the general and powerful tool [4–7], it is limited in practical engineering applications for its computational expenses. So, we consider the lowest-order quantum correction by taking only the $\alpha = 1$ term in (1), because the lowest-order term induces a major contribution in the quantum correction ingredients. Further, when the system is close to equilibrium, the approximate relations, $\nabla_r^3 U \simeq -k_B T \nabla_r \left(\nabla_r^2 \ln(n)\right)$ and $\nabla_k^3 f \simeq -(2\hbar^2/m^* k_B T)\nabla_k f$, are derived by using a displaced Maxwellian statistics, where n is the carrier density. By using these approximations, the following Boltzmann-like equation is obtained from (1) [2]

$$\frac{\partial f}{\partial t} + \frac{\hbar k}{m^*} \cdot \nabla_r f - \frac{1}{\hbar} \nabla_r \left(U - \frac{\hbar^2}{12m^*} \nabla_r^2 \ln(n)\right) \cdot \nabla_k f = \left(\frac{\partial f}{\partial t}\right)_C. \qquad (2)$$

The quantum effects are incorporated in terms of the quantum mechanically corrected potential in the driving term. Here, it is worth noting that the corrected term is identical to the quantum correction energy in the quantum moment theory [8]. Recently, we have demonstrated that the novel transport eqn (2) is sufficiently applicable for nonequilibrium quantum transport analysis and multi-dimensional practical use [2]. The validity of the quantum transport eqn (2) was checked by comparing the current–voltage characteristics and the carrier-density distributions of a single-tunneling barrier structure simulated based upon eqns (1) and (2) and further the classical Boltzmann equation [2]. In this paper, we will apply the Monte Carlo method to solve the quantum transport equation. Although the quantum correction term in (2) is represented under the effective-mass approximation, a formulation considering the full-band structure is also possible, as described in the Appendix.

Based upon (2), the velocity and force for particles during the free flight are respectively given by

$$\frac{dr}{dt} = \frac{\hbar k}{m^*} \qquad (3)$$

$$\frac{dk}{dt} = -\frac{1}{\hbar} \nabla_r \left(U - \frac{\hbar^2}{12m^*} \nabla_r^2 \ln(n)\right). \qquad (4)$$

The velocity eqn (3) is the same as that used in the standard Monte Carlo technique, but the force eqn (4) is modified in the quantum transport so that the particles evolve under the enforcement by the classical built-in potential plus the quantum-corrected potential. The quantum-corrected potential serves to soften the potential variations that the particles feel at the quantum regions where the potential and the carrier density change abruptly. In the equations of motion (3) and (4), we have described the carriers as though they have well-defined positions and momenta simultaneously. Although this seems to be inconsistent with the quantum mechanical uncertainty principle, we can interpret that the r and k obtained from (3) and (4) represent the centers of positions and momenta of traveling wavepackets, respectively. This situation could be the same as the quantum mechanical treatment of scattering events in the semi-classical Monte Carlo approaches.

3. Simulation results

To investigate the validity of our proposal, we applied (3) and (4) to a simple GaAs/AlGaAs single-tunneling barrier at 300 K. The barrier width and its energy height are 2.5 nm and 0.22 eV, respectively. The
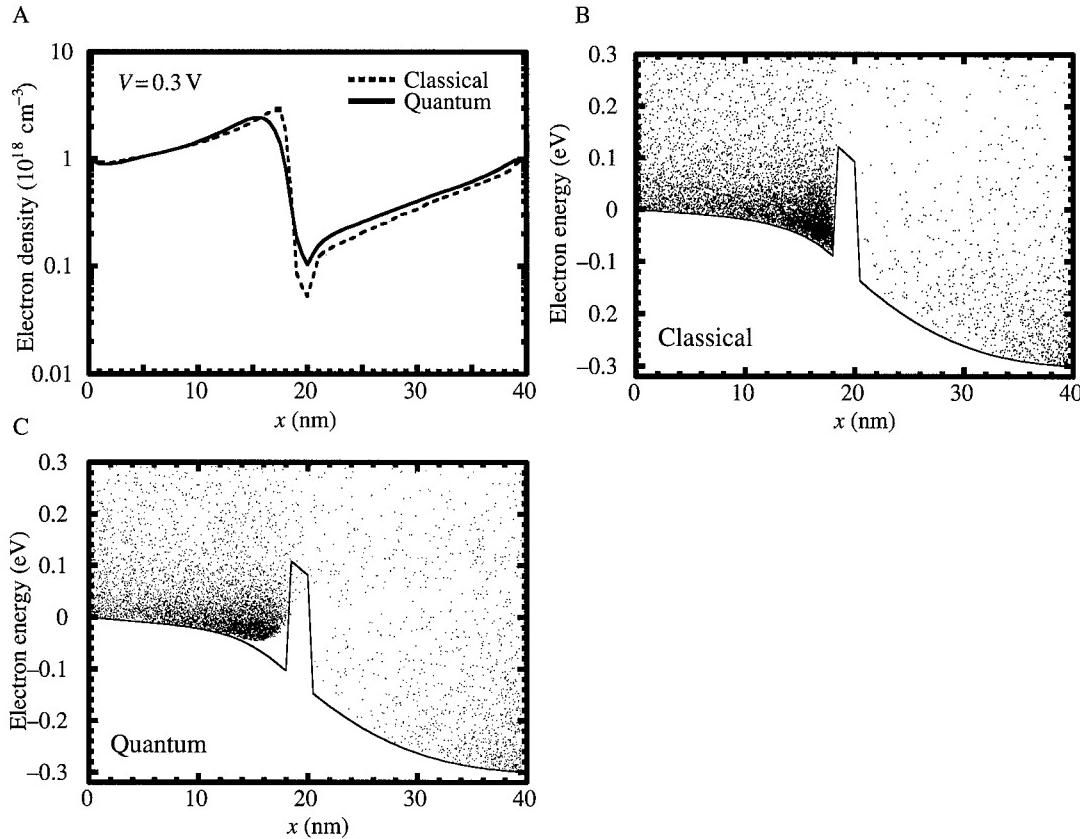
Fig. 1. Computed electron distributions of GaAs/AlGaAs single-tunneling barrier at 300 K, where the bias voltage is 0.3 V. A, indicates the electron density distributions, and B, and C, correspond to the particle distributions in space and energy for the classical and quantum Monte Carlo simulations, respectively. Here, for reference the conduction band profiles are also plotted by the solid line.

doping density in the GaAs electrodes is given as 10^{18} cm^{-3}. As a collisional process, the LO phonon scattering, the acoustic phonon scattering and the ionized impurity scattering are considered. The boundary conditions for the equations of motion (3) and (4) are given at the reservoirs attached to the quantum device as follows. The distribution of carriers injected into the device from a reservoir is characterized by the thermal equilibrium distribution of carriers at the reservoir, and all carriers impinging upon a reservoir from the device are absorbed by the reservoir. In addition, the carrier densities at the boundaries are determined to satisfy the charge neutrality condition. Figure 1 shows the computed electron distributions at the bias voltage of 0.3 V. Figure 1A indicates the electron density distributions, and B and C correspond to the particle distributions in space and energy for the classical and the quantum Monte Carlo simulations without and with the quantum corrected potential, respectively. For reference, the conduction band profiles are also plotted by the solid line. Here, please note that the vertical axis of Fig. 1C denotes the total electron energy including the quantum-corrected potential. Comparing Fig. 1B and C, we can observe the two major quantum effects in Fig.1C. The first one is the quantum confinement effects in the triangular potential well formed in the left electrode. The electrons are distributed away from the left barrier interface, and more surprisingly no electrons exist under a certain energy due to the quantized subband formed in the triangular potential well. The second one is the quantum tunneling effects through the central barrier. In the classical simulation of Fig. 1B, only the thermally excited electrons with energy larger than the barrier height are found in the barrier region. On

the other hand, in the quantum transport simulation of Fig. 1C a few tunneling electrons are detected apart from the thermally excited ones. This is due to the fact that the actual potential barrier height that electrons feel is lowered when the quantum corrected potential is considered. Here, the thermally excited electrons in the barrier region are found to be smaller in the classical simulation of Fig. 1B. We consider that this is because the classically simulated electrons are passing through the potential barrier more quickly, since a slightly higher electric field is induced at the barrier region due to the classical space charge distribution. Consequently, the quantum transport effects have been incorporated in the standard Monte Carlo techniques ingeniously by considering the quantum-corrected potential in the driving force.

4. Conclusion

In this paper, we propose a quantum Monte Carlo device simulation applicable for ultrasmall semiconductor devices, where the quantum effects are represented in terms of quantum mechanically corrected potential in the Boltzmann equation. In our model, the full-band structure of crystals can also be taken into account. We have demonstrated that the quantum tunneling and energy quantization effects of carriers can be incorporated in the standard Monte Carlo techniques successfully by using our model. We believe that the proposed quantum Monte Carlo technique will provide a powerful tool in the device simulation of ultrasmall MOSFETs in ULSI circuits.

References

[1] K. Hess, *Monte Carlo Device Simulation: Full Band and Beyond* (Kluwer Academic Publishers, 1991).
[2] H. Tsuchiya and T. Miyoshi, IEICE Trans. Elect. **E82-C**, 880 (1999).
[3] E. Wigner, Phys. Rev. **40**, 749 (1932).
[4] W. R. Frensley, Phys. Rev. **B36**, 1570 (1987).
[5] N. C. Kluksdahl, A. M. Kriman, and D. K. Ferry, Phys. Rev. **B39**, 7720 (1989).
[6] F. A. Buot and K. L. Jensen, Phys. Rev. **B42**, 9429 (1990).
[7] F. Rossi, C. Jacoboni, and M. Nedjalkov, Semicond. Sci. Technol. **9**, 934 (1994).
[8] J. R. Zhou and D. K. Ferry, IEEE Trans. Elect. Dev. **ED-40**, 421 (1993).

Appendix

When the full-band structure of crystals is considered, the third-order momentum derivative of the distribution function in the lowest-order correction term becomes $\nabla_k^3 f \simeq -(2\hbar^2/\mu(k)k_B T)\nabla_k f$, where $\mu(k)$ is defined by

$$\frac{1}{\mu(k)} = \frac{1}{2\hbar^2}\left[\nabla_k^2 E_k - \frac{1}{k_B T}|\nabla_k E_k|^2 + \frac{\nabla_k|\nabla_k E_k|^2 \cdot \nabla_k E_k}{|\nabla_k E_k|^2} - \frac{k_B T \nabla_k^3 E_k \cdot \nabla_k E_k}{|\nabla_k E_k|^2}\right], \qquad \text{(A.1)}$$

where E_k denotes the full-band energy structure. Consequently, the quantum transport equation for the full band model is represented as follows

$$\frac{\partial f}{\partial t} + v \cdot \nabla_r f - \frac{1}{\hbar}\nabla_r\left(U - \frac{\hbar^2}{12\mu(k)}\nabla_r^2 \ln(n)\right) \cdot \nabla_k f = \left(\frac{\partial f}{\partial t}\right)_C. \qquad \text{(A.2)}$$

As a matter of course, (A.2) reduces to the transport eqn (2) when the isotropic effective mass m^* is assumed.

Boundaries, transients and multiple barrier devices

H. L. GRUBIN, R. C. BUGGELN, J. P. KRESKOVSKY

Scientific Research Associates, Inc., P.O. Box 1058, Glastonbury, CT 06033, U.S.A.

Transient accurate numerical simulations of the Wigner function, with emphasis on boundaries, transients, barriers and dissipation are discussed. The absence of dc hysteresis and the presence of transient hysteresis when a double-barrier RTD is subjected to a controlled transient change in applied bias, highlights the nagging issue of the origins of hysteresis.

© 2000 Academic Press

Key words: Wigner functions, quantum transport, RTD, hysteresis.

The model and the approximations

A common feature in the semiconductor device industry is the use of numerical simulation in device design. Device design may, if required, incorporate physical models, analytical representations of physical models, the role of electrons and holes, band structure, transients and the external circuit. The situation does not change when we are dealing with quantum structures. Rather, the equations are more complicated, the issues deeper, and the physics less transparent. Additionally the reduced size, instead of decreasing the computational time for a simple current–voltage characteristic, is often orders of magnitude longer than for a comparable classical device. Additionally there is the nagging issue concerning the origins of hysteresis in the current versus voltage—is the hysteresis present under steady state conditions or is it a transient phenomena?

To address the above issues workers have developed and implemented algorithms for solving the Wigner distribution function, the density matrix (in a variety of different representations), nonequilibrium Green's functions, the Schrödinger equation, etc. In this study we focus on solutions to the time-dependent Wigner equation of motion. What is it we have done that is new?

To begin we only deal with devices whose lengths are long enough to incorporate flat-band contact regions. For these device simulations we solve the Wigner equation of motion with transients and dissipation. The dissipation model is closely coupled to the boundary conditions and is designed to reduce to the classical Ohm's law behavior in the absence of any quantum structure. The dissipation includes a regional relaxation time. For this study five distinct regions identify the relaxation model. The end regions (two) represent the contact regions, where the relaxation time takes on a very small value. (In the calculations below, we take the relaxation time to be 10^{-14} s, although some calculations were performed with values as low as 10^{-20} s.) The central region incorporates the quantum barriers and wells. Within this region the relaxation time is significantly larger, 10^{-12} s. In the transition regions (two) the relaxation time varies from its contact region value to the central region value in a continuous manner. In developing this model great care was taken to assure that the symmetry properties of the partial differential equation are mimicked in the difference equation [1]. To achieve this the spatial derivative was discretized at the half index spatial points using the second-order accurate central difference approximation. Alternatively, see [2–4].

The equilibrium distribution function is obtained from self-consistent solutions to the Wigner equation and Poisson's equation under zero-bias conditions without dissipation. In equilibrium we assume equilibrium Fermi–Dirac conditions on the boundary. Under nonequilibrium conditions, this boundary condition is replaced by one in which the normal spatial derivative of the Wigner function is zero. The consequences of this are flat-band conditions in the vicinity of the boundaries and a distribution that has the appearance of a *displaced Fermi distribution on the boundary*. We do not assume displaced Fermi conditions.

The potential distribution consists of the self-consistent Hartree potential energy obtained from Poisson's equation and the potential arising from the barriers and wells. Poisson's equation is solved throughout the entire device structure. The barriers and wells are square and their contributions to the Wigner integral are obtained analytically prior to any numerical computation. Presently, the contribution from Poisson's equation is incorporated classically through a derivative of the distribution function with respect to momentum. Under nonequilibrium conditions, when the dissipation model is successful and flat-band conditions occur, the fields at both ends of the structure are the same.

A variety of structures were studied. The one discussed below is 200 nm long with parameters appropriate to GaAs. The effective mass is constant throughout the structure. In the center of the structure are two 250 meV barriers, 5 nm wide, separated by 5 nm. The background doping was set to 10^{24} m^{-3} everywhere except for a 40 nm wide region in the center of the structure where the doping was 10^{21} m^{-3}. The transition from the low to high doping values occurred over a distance of 5 nm on either side of the barrier. The contact regions on either side of the structure are each 15 nm long, with the transition region 10 nm long.

The equation of motion of the Wigner function [5] as used in this study is:

$$
0 = \frac{\partial f_w(k, x)}{\partial t} + \frac{f_w(k, x) - f_{w0}(k, x)}{\tau(x)} + \frac{\hbar k_x}{m} \frac{\partial f_w(k, x)}{\partial x}
$$
$$
- \frac{1}{\pi \hbar} \lim_{L \to \infty} \int_{-L}^{+L} d\zeta \left[\begin{array}{c} V(x + \zeta) \\ -V(x - \zeta) \end{array} \right] \int dk'_x f_w(k'_x, k_y, k_z, x) \sin[2(k'_x - k_x)\zeta].
\tag{1}
$$

The dissipation model in eqn (1) does not, generally, yield local charge neutrality as:

$$
\int dk \left(\frac{f_w(k, \bar{\eta}) - f_{w0}(k, \bar{\eta})}{\tau(\bar{\eta})} \right) \neq 0.
\tag{2}
$$

However, in the vicinity of the contacts we obtain approximate charge neutrality.

The results

While the calculations of interest are transient ones we briefly discuss the dc characteristics, shown in Fig. 1. Two items are of interest. The first and most significant item is the absence of hysteresis for the converged solutions. Initially calculations were performed with voltage increments and decrements in units of $k_B T$. For this case we observed hysteresis in the current–voltage relation for values of voltage near the peak value of current. Upon closer examination, it was found that these solutions did not satisfy a pre-specified strict criterion of convergence. Introducing smaller changes in the applied voltage, the results shown in Fig. 1 were obtained. There remains a small voltage region, approximately $0.2\ k_B T$ near the peak current, where the dc computations need further study, as they did not satisfy the more stringent convergent criteria. The second item of issue is the structure of the Wigner function. At low and moderate values of bias the Wigner function is smooth, but picks up harmonic contributions as the peak current is approached, only to lose these contributions when the peak is past. These harmonics are reduced with smaller scattering time and with a combination of reduced grid size and increased momentum range.

However, the principle issue we are addressing is the determination of device speed. Placed in another context, if a device is subjected to an excitation, what is the shortest time interval before the device can be
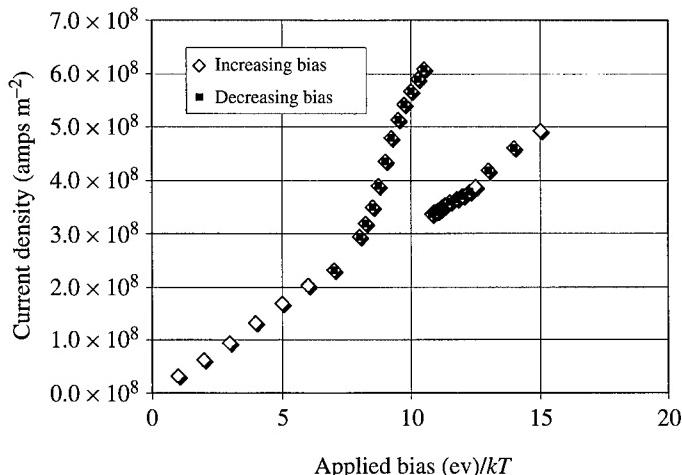
Fig. 1. dc current–voltage characteristic for the double-barrier diode.

interrogated about its new state? Or how long does it take the RTD to reach equilibrium. We address this problem by recognizing that an important RTD application is as an element in a circuit sustaining relaxation oscillations. In this simple context we may expect that the device to function as a clock at frequencies in excess of 100 GHz. Now negative differential conductance devices when part of an external circuit or transmission line, will under suitable conditions respond to pulse-like changes in applied bias with intra-pulse repetitive current–time and voltage–time profiles. This means that the only meaningful study of the response time of the device requires that we incorporate external circuit elements. But we should be able to address some minimal issues such as switching times without the elaborate use of the transmission line.

What we have done is to pulse the RTD starting from an initial steady state. In this calculation we include displacement as well as particle current contributions at the contact regions. (In the vicinity of the contacts the density is very nearly equal to the background density and so the total current at the boundaries is well represented by the sum of displacement and particle currents.) One calculation is illustrated in Fig. 2. For this case starting from a steady state in which the applied bias was $-8.0\ k_B T$, we dropped the potential energy at a controlled rate to $-12.0\ k_B T$, let the device sit at this value for approximately 700 fs, and then returned to the steady-state bias of $-8.0\ k_B T$.

In Fig. 2 we have superimposed two plots: the voltage pulse (trapezoidal shaped curve), and the total current at one boundary. (The time step in this calculation was 25 fs.) As the collector bias is made more negative with respect to the emitter, the current increases. The dominant contribution is displacement current. (Supplementary calculations demonstrate that the magnitude of this contribution increases with increasing rate of change of bias.) The subsequent relaxation to steady state is accompanied by positive and negative values of current, with the individual particle current contribution displaying opposite (signs) transient values. These opposite values indicate that in relaxing to steady states carriers can simultaneously either leave or enter the device through the emitter and collector contacts. After 800 fs the bias was changed to its initial value. The total current change is smaller, and at a given value of bias the transient current is different for the same, but rising and falling bias values. There is transient hysteresis!

For Fig. 2 calculations the system did not reach its steady state value at the end of 800 fs. There were still residual oscillations. While this residual oscillation is not likely to prevent the device from undergoing sustained oscillations with a period of the order of 700 fs, it may constrain the amplitude of the current oscillations when the device is an element in a circuit.
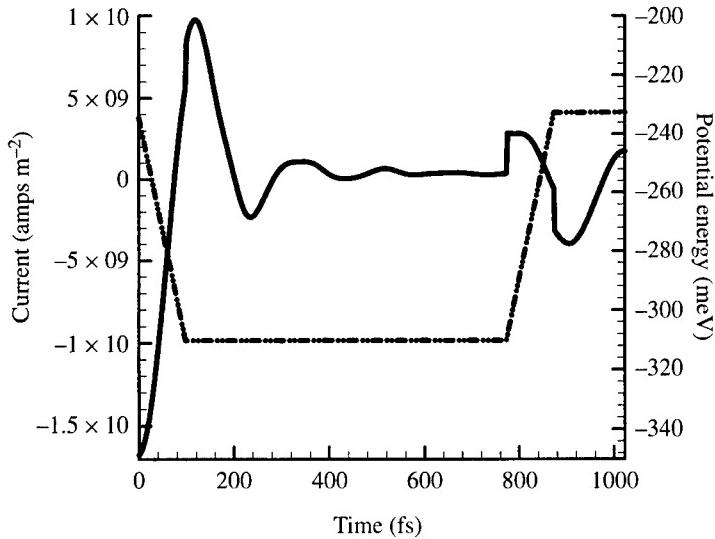
Fig. 2. Current density in response to a trapezoidal change in applied bias.
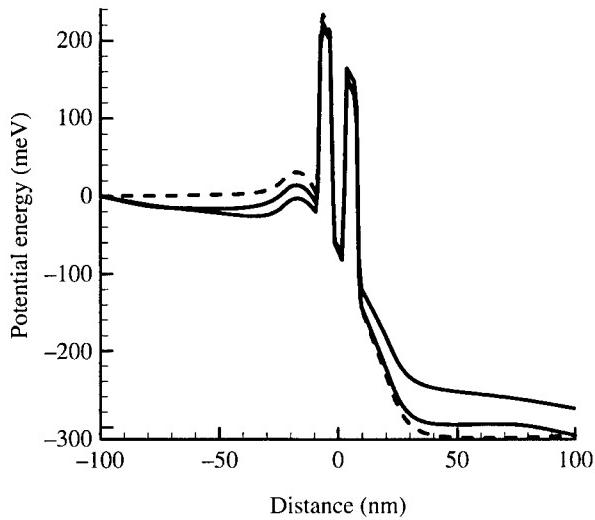
Fig. 3. Spatial distribution of potential energy at three different points in time.

It is important to note that other calculations in which the bias change was larger and moved well to the right of the voltage at peak current, $-15.0\ k_B T$, appeared to reach steady state at earlier times then the calculations displayed. Simply stated, the transient to steady state is bias dependent.

The charge distribution accompanying this current oscillation tends to follow, to a large extent, the steady state distribution. There is a build up of charge in the quantum well prior to the peak current and then an emptying of the well, beyond the peak current. Carriers simultaneously move into/out of the device accompanying the transient filling/emptying of the quantum well. This movement is reflected in the distribution of charge in the vicinity of the contacts, the subsequent potential distribution and the subsequent displacement

current. However, this change is continuous in time, and to illustrate, we display in Fig. 3 the potential distribution within the device at three points in time. Note that the change in slope and in the sign of the slope of the potential distribution manifests itself in displacement current contributions.

Details of the charge distribution show some familiar contact–region boundary contributions. For example, in the calculation above there was a smaller scattering time within the contact region. With the collector negative with respect to the emitter there was a region of local charge accumulation in the emitter region arising solely from the change in scattering time. On the collector side there was a corresponding local depletion of carriers. The amount of accumulation/depletion is bias dependent (and current dependent) with neither accumulation nor depletion under equilibrium conditions. These results were expected and are a common occurrence in classical devices with similar scattering variations.

Conclusions

What does this all mean? First and most important the results indicate that the initial transients of these quantum structures are dominantly capacitive and associated with displacement current contributions. Second, the derivative boundary condition on the distribution function does not constrain us to deal with displaced boundary distributions. As such we may be able to deal with more realistic contact conditions. Third, the absence of hysteresis for the dc characteristic indicates that such phenomenon is primarily transient, as in classical devices. But it must be cautioned that we have not done an exhaustive study of the observed hysteresis and so any conclusion regarding dc hysteresis must necessarily be regarded as tentative.

References

[1] N. C. Kluksdahl *et al.*, Self-consistent study of the resonant-tunneling diode, Phys. Rev. **B39**, 7720 (1989).

[2] B. A. Biegel, *Quantum Electronic Device Simulation.* Ph.D. dissertation submitted to Department of Electrical Engineering at Stanford University (1997).

[3] W. R. Frensley, Boundary conditions for open systems driven far from equilibrium, Rev. Mod. Phys. **62**, 745 (1990).

[4] K. L. Jensen and F. A. Buot, The effects of scattering on current–voltage characteristics, transient response, and particle trajectories in the numerical simulation of resonant tunneling diodes, J. Appl. Phys. **67**, 7602 (1990).

[5] E. Wigner, On the quantum correction for thermodynamic equilibrium, Phys. Rev. **40**, 749 (1932).

Computational electromagnetic methods for interconnects and small structures

CONSTANTINE A. BALANIS, ANASTASIS C. POLYCARPOU, STAVROS V. GEORGAKOPOULOS

Department of Electrical Engineering, Telecommunications Research Center, Arizona State University, Tempe, AZ 85287-7206, U.S.A.

The continual advances in speed and integration scale of electronic circuits have created enormous demands for high-speed, high-density packages which ensure reduced interconnection delays and improved electrical performance. Such structures usually involve a large number of planar transmission lines at various levels within the package, whereas the geometrical orientation of these lines is not necessarily uniform. Also, the existence of multiple dielectric layers, discontinuities, bends, and wire bounds adds considerable complexity to the package. It is therefore essential that full-wave computational electromagnetic (CEM) techniques, such as the finite element method (FEM) and the finite-difference time-domain (FDTD) method, be developed and used to accurately model the electrical performance of these devices and circuits.

© 2000 Academic Press

Key words: computational electromagnetic methods, interconnects, small structures.

The finite element method (FEM) and the finite-difference time-domain (FDTD) are used here to model three-dimensional circuits and packages that are frequently found in high-frequency applications. Numerical predictions are often compared with results obtained from the literature.

A low-pass filter, originally examined by Sheen *et al.* [1], is shown in Fig. 1. The microstrip lines are printed on a Duroid substrate of $\epsilon_r = 2.2$ and height 0.794 mm. This geometry was simulated using the in-house developed FEM and FDTD codes for a frequency span of 20 GHz. The magnitude of S_{21} vs. frequency is illustrated in Fig. 1 which illustrates an excellent agreement between the two methods. The CPU time required for the FEM simulation is on average 15 min per frequency point. This problem was simulated on a 370 IBM RISC/6000 UNIX workstation. The corresponding CPU time for the FDTD simulation is a total of 45 min.

The interconnection of two microstrip lines, which are printed on separate substrates (GaAs), is analysed using the FEM. The microstrip transition between the two modules is encapsulated by a dielectric region (bridge) with relative dielectric constant ϵ_r and length d. The geometry of this structure, which was analysed prior to this work using the FDFD method [2], is shown in Fig. 2. The sidewalls are perfect electric conductors since the structure is shielded. The magnitude of S_{11} vs. the normalized frequency $(k_0 \cdot a)$ is plotted for various values of ϵ_r and d. The dimension a used in the normalization was taken to be 1 mm whereas the frequency
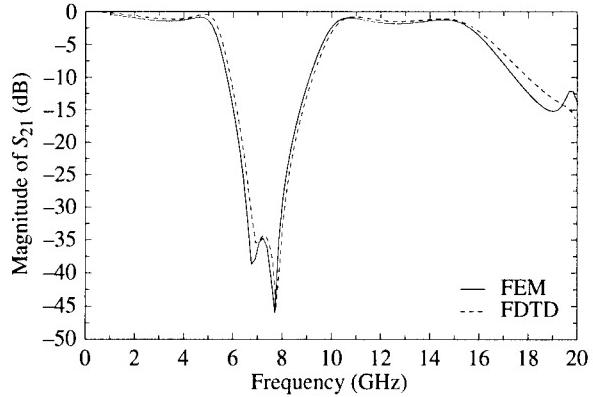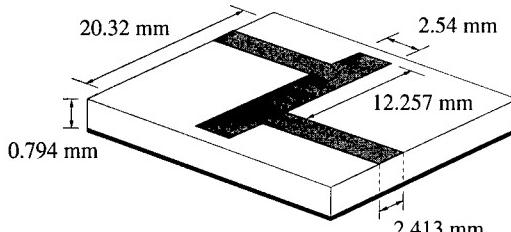
Fig. 1. Low-pass filter on a Duroid substrate of $\epsilon_r = 2.2$.
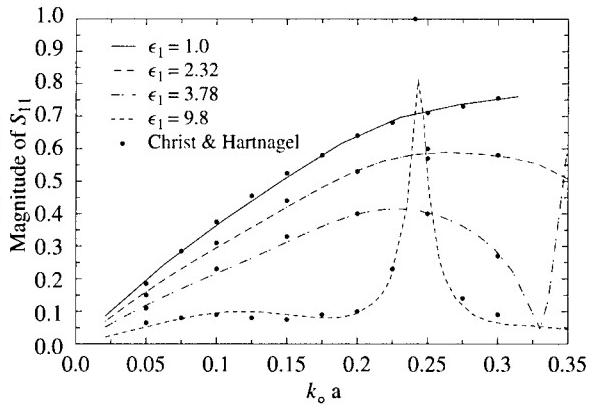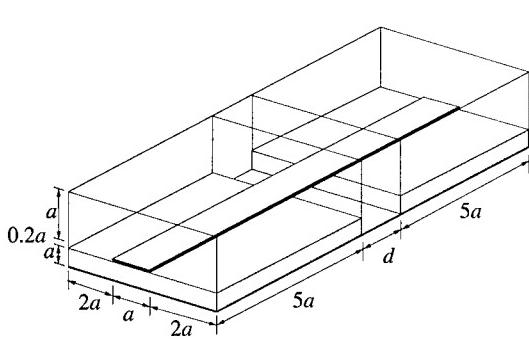
Fig. 2. Microstrip transition through a dielectric bridge ($d = 3.17a$).

of operation ranges between 100 MHz and 15 GHz. By increasing the permittivity of the dielectric bridge, it is expected that the microstrip transition will provide a better match since the effective dielectric constant (ϵ_{reff}) of the microstrip line on a GaAs substrate is between 7.54 and 9.81. Thus, for a dielectric bridge with dielectric constant close to ϵ_{reff}, the microstrip transition should result in substantially less reflections compared with the air bridge. The dependence of $|S_{11}|$ on the dielectric constant of the bridge is illustrated for $d = 3.17a$ in Fig. 2. As shown, a dielectric bridge with $\epsilon_r = 9.8$ provides the best match of all four cases for a normalized frequency of $k_0 \cdot a < 0.2$. At higher frequencies, the dielectric bridge, which acts like a resonator, causes the magnitude of S_{11} to rapidly increase.

The geometry shown in Fig. 3, which was previously analysed in [2], represents a wire bond between two microstrips printed on separate substrates. The substrate material is GaAs with $\epsilon_r = 12.9$. The structure is placed inside a waveguide with transverse dimensions $5a \times 3.2a$. The bond wire is rectangular in shape with cross-sectional area $0.2a \times 0.2a$ and an inner length d. All conducting surfaces were considered as being perfect electric conductors. The normalization constant a was 1 mm, whereas the normalized frequency range is given by $0 < k_0a \leq 0.3a$. The magnitude of S_{11} for this structure is plotted in Fig. 3 for three different lengths d. From this figure, it is evident that there is an excellent agreement between our results and results
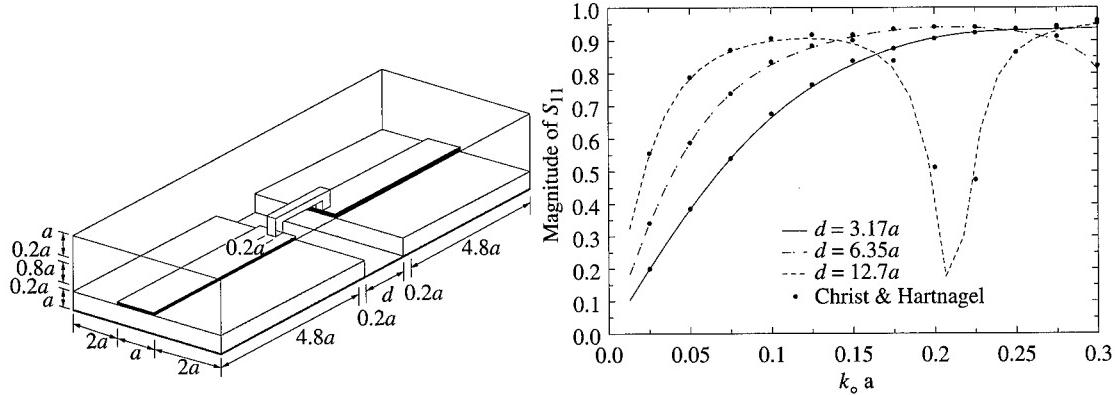
Fig. 3. Bond-wire interconnection of two microstrip lines.
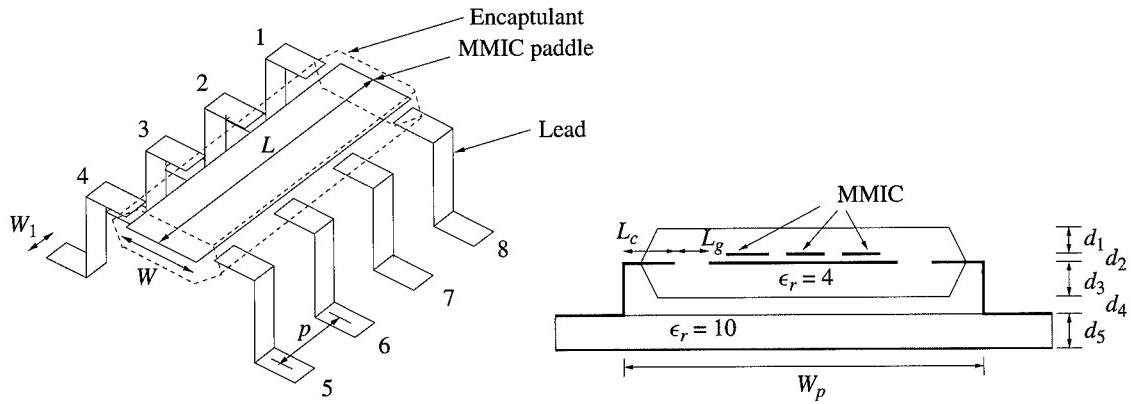
Fig. 4. Geometry of SOIC-8 plastic package: $L = 4.2$ mm, $W = 2.4$ mm, $p = 1.27$ mm, $w_1 = 0.4$ mm, $W_p = 4.4$ mm, $L_c = 0.7$ mm, $L_g = 0.3$ mm, $d_1 = 0.535$ mm, $d_2 = 0.1$ mm, $d_3 = 0.635$ mm, $d_4 = 0.2$ mm, $d_s = 0.25$ mm.

obtained using the FDFD method [2]. An increase in the length of the bond wire results in a larger overall S_{11}. At higher frequencies, the structure resonates thereby forcing S_{11} to zero. However, compared with the microstrip interconnection through an air bridge, which is shown in Fig. 2, the bond wire creates higher reflections. For example, at a frequency of 4 GHz, a bond wire with length 3.17 mm and height 0.2 mm causes more than 50% reflections. On the contrary, the microstrip interconnection through an air bridge, at the same frequency, causes only 25% reflections.

The electrical performance of an MMIC is drastically altered when mounted into a package such as the eight-lead small outline integrated circuit (SOIC-8) surface-mount plastic package shown in Fig. 4. Plastic packages have been used for years up to a frequency of about 2.5 GHz. It becomes extremely difficult to utilize these packages for higher frequencies mainly because of poor grounding of the paddle and package resonances. First, the MMIC shown in Fig. 5 is analysed in the absence of the package. This circuit represents a basic through-connection with a stub. When placed into the package, the circuit itself is elevated 0.1 mm above the ground and embedded into a dielectric medium with dielectric constant $\epsilon_r = 4$. The same MMIC structure, mounted into the SOIC-8 surface-mount package, was analysed by Jackson [3] using the method
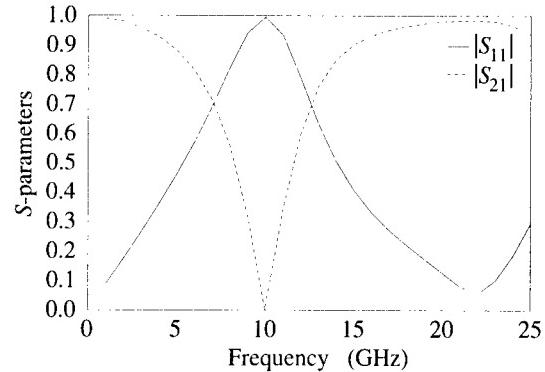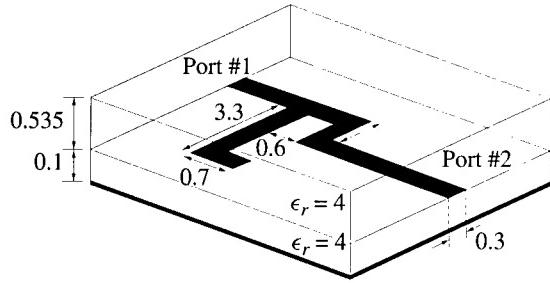
Fig. 5. Geometry of the unpackaged MMIC and corresponding S-parameters.
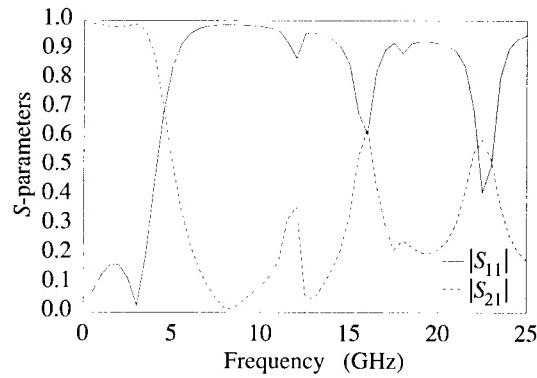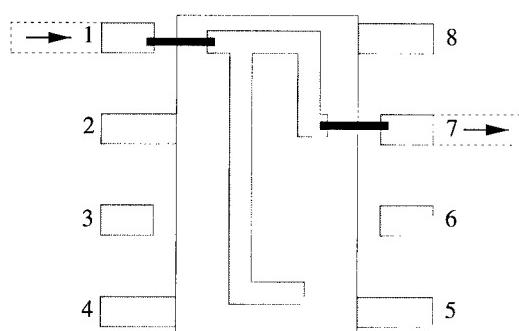
Fig. 6. Top view of the SOIC-8 plastic package and corresponding S-parameters.

of moments (MoM). The S-parameters of the through-connection were calculated using the FEM. The magnitude of S_{11} and S_{21} is shown vs. frequency in Fig. 5. It is evident that the circuit resonates at 10 GHz which corresponds to the frequency where the stub becomes quarter of a wavelength long.

The electrical performance of the MMIC is totally altered when mounted into the SIOC-8 plastic package which is shown in Fig. 4. The input port of the circuit is connected through bond wire of rectangular cross-section (0.1 mm × 0.1 mm) to lead 1, and the output port of the circuit is connected through an identical bond wire to lead 7, as shown in Fig. 6. The height of the bond wire, from the top surface of the lead or paddle to the inner surface of the horizontal wire, is 0.35 mm whereas the bond-wire length, as measured from the inner surfaces of the vertical vias, is 0.6 mm. Leads 2, 4, 5 and 8 are directly connected to the paddle and also grounded through vertical conducting vias to the motherboard. Leads 3 and 6 are grounded to the motherboard but not otherwise connected to the paddle. Figure 6 shows the corresponding S-parameters of the entire package. Compared with the unpacked structure, whose S-parameters are illustrated in Fig. 5, the first resonance shifts from 10 to 8 GHz. The downshift of the resonant frequency of the MMIC is due to distracting interference between the circuit and the package. Additional package resonances are also observed in Fig. 6. These correspond to 12, 15.5 and 22.5 GHz.

References

[1] D. M. Sheen, S. M. Ali, M. D. Abouzahara, and J. A. Kong, Application of the three-dimensional finite-difference time-domain method to the analysis of planar microstrip circuits, IEEE Trans. Microw. Theory Tech. **MTT-38**, 849 (1990).

[2] A. Christ and L. Hartnagel, Three-dimensional finite-difference method for the anlaysis of microwave-device embedding, IEEE Trans. Microw. Theory Tech. **MTT-35**, 688 (1987).

[3] R. W. Jackson, A circuit topology of microwave modeling of plastic surface mount packages, IEEE Trans. Microw. Theory Tech. **MTT-44**, 1140 (1996).

Ionic channels: natural nanotubes described by the drift diffusion equations

BOB EISENBERG[†]

Department of Molecular Biophysics and Physiology, Rush Medical College, 1750 West Harrison, Chicago, IL 60612, U.S.A.

Ionic channels are a large class of proteins with holes down their middle that control a wide range of cellular functions important in health and disease. Ionic channels can be analysed using a combination of the Poisson and drift diffusion equations familiar from computational electronics because their behavior is dominated by the electrical properties of their simple structure.

© 2000 Academic Press

Key words: tonic channels, natural nanotubes, drift diffusion equations.

1. Generalities

Mesoscopic devices are relatively new to engineering but—in the form of proteins—they have performed a surprising fraction of the work of life for a very long time. Proteins are the preferred device of evolution, because evolution acts directly on genes, which can only make proteins. Genes are the one-dimensional blueprints for proteins. Genes specify the amino acids (peptides) which are linked like beads into a string to make a protein. Each bead has different fixed charge and (somewhat) different chemical properties. The polypeptide string becomes the functioning protein when it is folded into a three-dimensional structure that can act as a device.

Many proteins should be studied as devices. Most proteins function far from equilibrium and have well-defined inputs and outputs. Neither the structure nor the chemistry of most proteins make much sense if studied at equilibrium and thus isolated from function. Proteins at equilibrium are dead, about as interesting as a transistor with its terminals soldered together. Ionic channels are a class of proteins that have been studied as devices with some success.

2. Specifics: ion channels

Ionic channels control the flow of substances in and out of cells and, by controlling the flow of ions, they are produce nearly all the electrical activity of living systems. A substantial fraction of the drugs used by physicians act directly or indirectly on channels.

Channels come in many different types. Hundreds of types are now known [1], and thousands probably remain to be discovered. Each channel type has a specific role in a biological system and is (usually) well

[†]E-mail: beisenbe@rush.edu

tuned for its task. Some types of channels are devices that respond to mechanical force ('touch'). Others respond to electrical potential (and thereby create the action potentials that are the digital signals of our nervous system). Many respond to specific chemicals. Channels are devices with specialized inputs that control specific outputs, the flow of particular types of ions across cell membranes. Channels are highly selective in how they respond and what they respond to.

Ionic channels function in two ways. They open and close in response to stimuli in a process called gating. Once open, they select between different species of ions. (Remember that in biology current is nearly always carried by ions: Na^+, K^+, Ca^{++}, and Cl^- are the chief carriers of current and each is controlled by different types of channel proteins. Indeed, there are tens to hundreds of types of K^+ channels, Ca^{++} channels and so forth.)

Unfortunately, not enough is yet known to analyse gating as a property of the device, because we do not know the structure or basic operating mechanism of a channel with typical gating. (The structure of only a few types of channels are known and as of today they do not include one with typical gating properties.) This situation is likely to change dramatically in the near future, but as of now, device analysis works best on the other main property of channels, its conduction of current. The structure of a channel does not change (in the mean) while it conducts current. An ionic channel is a hole in the wall that functions by electrodiffusion. If there is any protein that we should be able to analyse as a device, it is an ionic channel.

3. Traditional kinetic models

Open channels have until recently mostly been studied in the tradition of gas phase chemical kinetics [2] but the chemical tradition has not been successful in predicting their function [3], namely in predicting the current that flows through them under a range of conditions, because gases are so different from the condensed phases of liquids and proteins.

Simulations of the molecular dynamics of proteins and channels are appealing and popular [4, 5] but their high-resolution produces serious limitations in size and scope. Molecular dynamics done in the biological tradition has not included a definite concentration of ions, nor a potential across a channel, nor has it been able to compute current through the channel. Thus, it has little to say about the channel as a device despite the power with which it describes a channel as a structure. These limitations may be removed if the Monte Carlo methods of computational electronics are used to simulate channels.

Channel proteins can be described at lower resolution as a distribution of fixed charge using a mean field theory nearly identical to the drift diffusion theory of computation electronics. Extensive experimental evidence supports the theory [6]: it fits data from some seven types of channels measured over a wide range of voltages and concentrations. Its parameters are reasonable and can be checked against independent experimental evidence in favorable cases.

The one-dimensional theory we use to describe an open channel represents the structure of the channel's pore as a cylinder of variable cross-sectional area $A(x)(\text{cm}^2)$ along the reaction path x (cm) with dielectric coefficient $\varepsilon(x)$ and a density of charge $\rho(x)(\text{coul} \cong \text{cm}^{-1})$. eN_A is the charge in 1 mole of elementary charges e, i.e. the charge in a Faraday. The charge $\rho(x)$ is as follows.

(1) The charge $eN_A \sum_k z_k C_k(x)$ of the ions (that can diffuse) in the channel, of species k of charge z_k, and mean concentration ; $C_k(x)$; typically $k = Na^+$, K^+, Ca^{++}, or Cl^{Γ}.

(2) The permanent charge of the protein $P(x)(\text{mol} \cong \text{cm}^{-1})$, which is a permanent part of the atoms of the channel protein (i.e. independent of the strength of the electric field at x) and does not depend on the concentration of ions, etc, and so is often called the fixed charge. $P(x)$ is really quite large because the channel is so small. One charge in a cylinder 6 Å diameter and 10 Å long is a concentration of 6×10^{21} cm$^{-3} \approx 10$ M. Interestingly, mean field theories in electrochemistry are known to 'become exact for large electric fields, independent of the density of hard spheres' [7, p. 315], 'independent of

interactions of molecules in the fluid phase' [8, p. 972]. Also, some channels are thought to have as many as six charges in half that length or volume, giving $\sim$100 M charge: NaCl in the selectivity filter of such channel may be more like table salt than sea water.

(3) The dielectric charge (i.e. the induced charge which is strictly proportional to the local electric field) is not included in $\rho(x)$ because it is described by $\varepsilon(x)$. It is generally small compared with the structural charge.

We make the usual mean field assumptions that the average charge $\rho(x)$ produces an average potential $\varphi(x)$ according to Poisson = s equation and that the mean electric field $-\nabla\varphi$ captures the properties of the fluctuating electric field which are important on the slow timescale of biology. These assumptions are hardly novel; indeed, it requires some extraordinary circumstances for them not to be true, in slow highly averaged systems

$$\varepsilon_0\left[\varepsilon(x)\frac{d^2\varphi}{dx^2} + \left(\frac{d\varepsilon(x)}{dx} + \varepsilon(x)\frac{d}{dx}[\ln A(x)]\right)\frac{d\varphi}{dx}\right] = -\rho(x) = -eN_A\left[P(x) + \sum_k z_k C_k(x)\right]. \quad (1)$$

The boundary conditions for the potential in the real world are set by the potentials in the baths surrounding the channel, i.e. the potential on the left is known and maintained at V_{appl} and that on the right is held at zero. The flow (i.e. flux J_k of ion k) through the channel is described in mean field theory by the diffusion equation, the Nernst–Planck equation (see [9, 10]; derived below) written in general form for channels of arbitrary variation in cross-sectional area $A(x)$ and diffusion coefficient $D(x)$

$$J_k = -D_k(x)A(x)\left(\frac{dC_k(x)}{dx} + \frac{C_k(x)}{RT}\frac{d}{dx}[z_k F\varphi(x) + \mu_k^0(x)]\right);$$

$$I = \sum_k I_k = \sum_k z_k F J_k. \quad (2)$$

The flux J_k of ions is driven by the (gradient of) concentration and electrical potential, which together form the electrochemical potential $\mu_k = RT\log_e C_k(x) + z_k F\varphi(x)$. $D_k(x)$ is the diffusion coefficient of ion k in the channel's pore.

Specific chemical interactions are important when dealing with selectivity in mixtures of ions [11–13] and can be described by an excess chemical potential $\mu_k^0(x)$. In the one case considered in detail up to now, $\mu_k^0(x)$ can be computed from the volume of the ions and charged groups of the protein. Additional chemistry is not needed to explain the selectivity of the L-type Ca channels of cardiac muscle. This is surprising because the L-type Ca channels distinguish between ions with remarkable selectivity.

The L-type Ca channel is made of the selectivity oxygens of the glutamate locus of the protein. The selectivity oxygens are described as tethered ions with the same properties as carboxylate ions in bulk but confined to the subvolume of the selectivity filter. Ions such as Ca^{++}, Na^+, and Cl^- can move from phase to phase, but the selectivity oxygens cannot. Ions bind in, or are excluded from, the filter because the system has a more (or less) favorable free energy when ions are bound than when they are free. The free energy of binding/exclusion involves 'excess' terms that arise from the finite volume of ions.

The novel part of the analysis is the *ab initio* computation of the thermodynamic excess properties of ions in the selectivity filter using a statistical mechanical theory of bulk electrolyte solutions, the so-called 'primitive' version of the mean spherical approximation MSA [14, 15]. This theory represents ions as charged hard spheres and water as a continuous dielectric. The mutual exclusion of the finite ionic volumes and the electrostatic interactions (screening) among the ions produce the nonideal ('excess') components of the chemical potentials. The excess chemical potentials are generally different for different ionic species. No other effects (such as specific interactions between atomic orbitals of Ca^{++} and the molecular orbitals of the carboxylic groups) are considered in this model of the selectivity filter.

Predictions of current through the channel agree with experiment over the entire range of Ca^{++} concentration from submicromolar to 100 mM, in the presence of 100 mM NaCl, if the selectivity filter of the channel has a dielectric constant of 80 and a volume of $\sim$375 $(nm)^3$ the model, and the ions and carboxyl oxygens are given their crystal radii.

L-type Ca channels distinguish between Na^+ and Ca^{++} in two ways: (1) Divalent Ca^{++} screens the carboxylate groups of the EEEE locus more effectively than monovalent Na^+, and thus has more negative and attractive electrostatic energy there. (2) Four Na^+ displace about twice the volume of two Ca^{++}, with a significant change in the volume fraction occupied by ions and a resulting more attractive excess energy for Ca^{++}. With electroneutrality as the dominating constraint, the Ca^{++} channel can use volume exclusion to distinguish ions of nearly the same diameter, such as Na^+ and Ca^{++}, because equal charge excludes different volumes depending on the valency of the permeating ion. In this way, electrostatics facilitates selection by volume exclusion: substantial repulsion due to excluded volume selects against Na^+ (and other monovalent cations) compared with Ca^{++}.

4. Conclusion

Even the most biological property of the open channel—its selectivity between ions—can be understood from a primitive model of the physical interactions of ions of finite volume. It seems likely that Monte Carlo simulations in the tradition of computational electronics can be applied productively to a wide range of biological phenomena if they describe ions as objects (usually spheres) of finite volume.

References

[1] F. M. Ashcroft, *Ion Channels and Disease* (Academic Press, New York, 1999) p. 481.

[2] B. Hille, *Ionic Channels of Excitable Membranes* (Sinauer Associates Inc, Sunderland, 1992) 2nd edn, pp. 1–607.

[3] C. Miller, Ionic hopping defended, J. Gen. Physiol. **113**, 783 (1999).

[4] C. L. Brooks, M. Karplus, and B. M. Pettitt, *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics* (John Wiley & Sons, New York, 1988).

[5] B. Roux and M. Karplus, Molecular dynamics simulations of the gramicidin channel, Ann. Rev. Biophys. Biomol. Struct. **23**, 731 (1994).

[6] B. Eisenberg, Ionic channels in biological membranes. Electrostatic analysis of a natural nanotube, Contemporary Physics **39**, 447 (1998).

[7] D. Henderson, L. Blum, and J. L. Lebowitz, An exact formula for the contact value of the density profile of a system of charged hard spheres near a charged wall, J. Electronal. Chem. **102**, 315 (1979).

[8] L. Blum, Contact theorems for rough interfaces, J. Stat. Phys. **75**, 971 (1994).

[9] J. Bockris and A. M. E. Reddy, *Modern Electrochemistry* (Plenum Press, New York, 1970) p. 1432.

[10] J. S. Newman, *Electrochemical Systems* (Prentice-Hall, Englewood Cliffs, NJ, 1991) 2nd edn, p. 560.

[11] W. Nonner, D. P. Chen, and B. Eisenberg, Anomalous mole fraction effect, Electrostatics, and binding in ionic channels, Biophys. J. **74**, 2327 (1998).

[12] W. Nonner and B. Eisenberg, Ion permeation and glutamate residues linked by Poisson-Nernst-Planck theory in L-type calcium channels, Biophys. J. **75**, 1287 (1998).

[13] D. Chen, *et al.*, Selectivity of calcium release channel (CRC) of cardiac muscle, Biophys. J. **74**, A342 (1998).

[14] J.-P. Simonin and L. Blum, Departures from ideality in pure ionic solutions using the mean spherical approximation, J. Chem. Soc. Faraday Trans. **92**, 1533 (1996).

[15] L. Blum, Mean spherical model for asymmetric electrolytes I: method of solution, Mol. Phys. **30**, 1529 (1975).

Quantum-dot transport in carbon nanotubes

T. IDA, K. ISHIBASHI[†], K. TSUKAGOSHI[‡], Y. AOYAGI[†]

The Institute of Physical and Chemical Research (RIKEN), 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan

B. W. ALPHENAAR

Hitachi Cambridge Laboratory, Madingley Road, Cambridge CB3 0HE, U.K.

Transport measurements on a bundle of single-walled carbon nanotubes have been made below 4.2 K as a function of side gate and source–drain bias voltage. The transport of an individual nanotube is described by the Coulomb blockade effect. The zero-dimensional quantum states of the nanotube become clear for measurements of large bias voltage. In addition, we present preliminary results of microwave application to the SWNT dot, and the results can be qualitatively explained by classical coupling to the dot.

© 2000 Academic Press

Key words: carbon nanotubes, Coulomb blockade, microwave irradiation effect, quantum dot.

Since the discovery of the graphite needles which are called carbon nanotubes [1] with cylindrical diameter of the order of 1 nm, much investigation has been done owing to their unique electronic and geometric properties. In particular, single-walled carbon nanotubes (SWNTs) are quasi one-dimensional quantum wires, and could be a building block of quantum nanostructures, showing quantum effects at higher temperatures. Recently, electrical transport of an individual carbon nanotube has been measured with attention paid to unique quantum effects [2–5].

For future quantum devices, in which quantum states would be manipulated, these effects have to be clearly resolved at reasonably high temperatures. However, in quantum dots made in semiconductor material with standard electron-beam lithography, the size is limited to several tens of nanometers. As a result, the Coulomb blockade effect and the zero-dimensional (0D) effect are observable typically below liquid-helium temperatures [6]. SWNTs could be interesting, in this respect, because they could be used for the element of very small quantum dots. In this report, we study basic transport characteristics of SWNTs and characterize them as quantum dots. The interaction between high-frequency fields and quantum dots is important for the coherent control of quantum states in future devices. As a first step, we present preliminary experimental results on microwave irradiation of SWNT quantum dots.

Figure 1A shows a schematic circuit of the quantum dot in a SWNT rope. The SWNTs had a diameter

---

© 2000 Academic Press
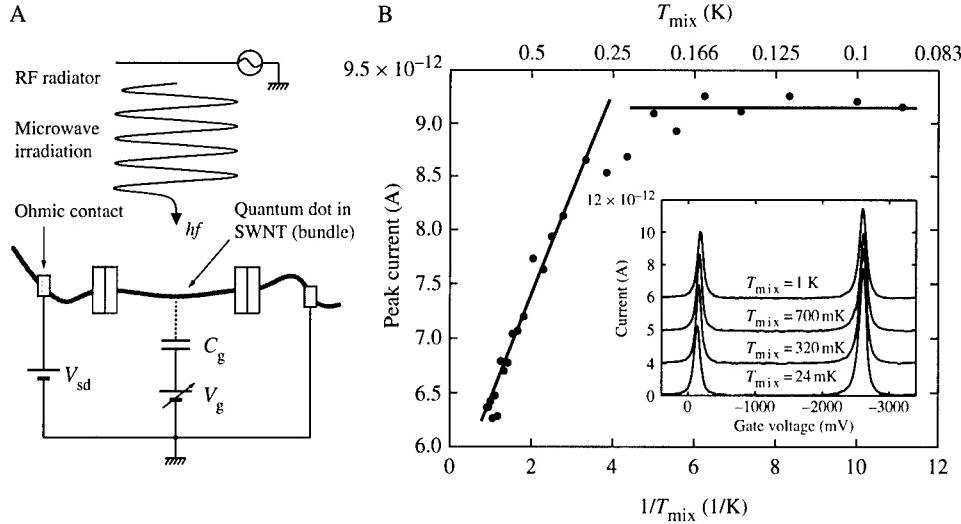
Fig. 1. A, Schematic measurement set-up of a quantum dot in SWNT. A constant bias voltage V_{sd} is applied to the SWNT. The tunnel barrier may exist at the junction of the electrode and SWNT. There is one side gate pad near the SWNT electrostatically coupled to the dot with capacitance C_g. The other side gate is used for the microwave irradiation. B, Temperature dependence of the Coulomb peak current. Inset: Coulomb blockade oscillations of SWNT at increasing temperatures.

ranging from 1 to 15 nm and a length of about 20 μm. These bundles were dispersed on an oxidized silicon surface. Then, two platinum/gold leads with 6 nm/50 nm thickness were patterned on one of the nanotubes by using the mark alignment technique in conventional electron-beam lithography. The distance between lead contacts was 250 nm. Two side gate pads near the nanotube between two contacts were also formed. One gate was used for varying the potential in the SWNT, while the other was used to apply the microwave field. The sample was mounted on the mixing chamber of the dilution refrigerator with a base temperature of 25 mK, and two-terminal dc measurements were performed. The $I-V_{sd}$ characteristic of the bundle was linear at room temperature with the resistance from 50 k to 100 kΩ, but became nonlinear at low temperature.

The inset in Fig. 1B shows the measured current as a function of the applied gate voltage V_g at various temperatures. The curves are offset for clarity. Two Coulomb blockade peaks are observable in the swept gate voltage range from 0 to -4 V. The basic behavior of Coulomb blockade oscillations is similar for all temperatures. The oscillations were very stable for each measurement, suggesting that no unstable charge states was around the dot. In Fig. 1B, we show the peak current as a function of the inverse of mixing chamber temperature, $1/T_{mix}$. At high temperatures, the peak current monotonically increases as T_{mix} is decreased, and the peak current saturation is observed at temperatures below $T_{mix} = 250$ mK. This indicates that the electron temperature stays constant even through the mixing chamber temperature further decreases.

Figure 2A shows the $I-V_{sd}$ characteristics for varying gate voltages from -1 to -6 V at $T_{mix} = 4.2$ K. The corresponding differential conductance is shown in Fig. 2B. In the figure, the Coulomb gap with a diamond shape is clearly observed. As seen in these figures, the measured device shows the typical characteristics of a single quantum dot. The lines seen outside of the diamond region in the gray-scale plot are due to the effect of 0D excited states, first observed in a GaAs quantum dot [7]. This effect becomes more pronounced at the lowest temperatures. We do not know how the tunneling barriers are formed in the device. Actually, we have measured a few other devices made in the same way, and all devices have shown similar behavior. This fact means that the tunneling barrier may not be due to some defects in nanotube but may be formed at the junction between the edge of the electrode and the nanotube.

The self-capacitance and the corresponding charging energy obtained from Fig. 2 are $C_\Sigma = 18.6$ aF and $E_C = 8.6$ meV, respectively, which are not easy to realize in a conventionally made GaAs dot. We estimate
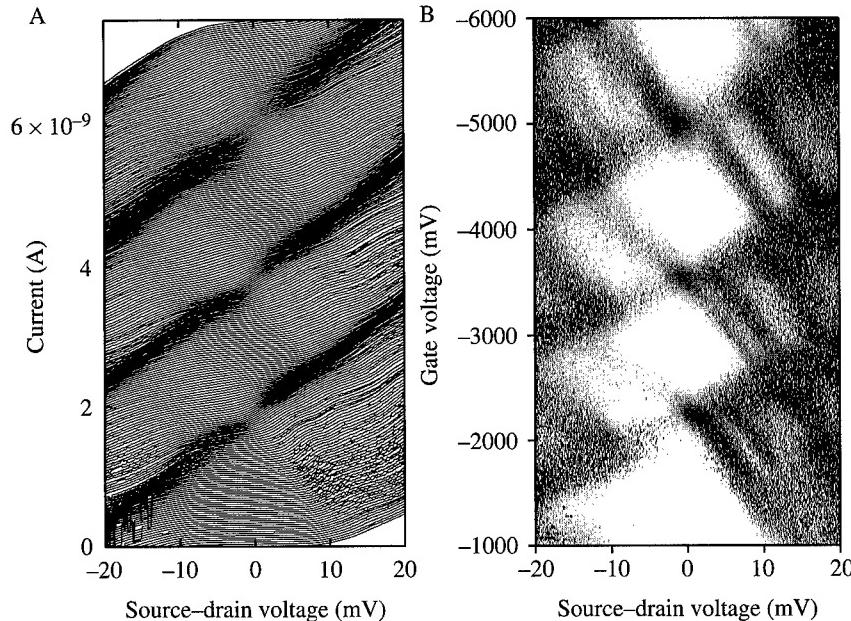
Fig. 2. A, I–V_{sd} characteristics at 4.2 K for gate voltage V_g ranging from -1 to -6 V with an increment of 20 mV. The curves are offset for clarity. B, Gray-scale plot of the differential conductance.

the energy level spacing $\Delta E = 2.3$ meV, which happens to be the similar value to E_C, from the measurements of Coulomb blockade oscillation for large source–drain bias voltages. The charging energy corresponds to a temperature of about 100 K. By narrowing the distance between the two contacts, we may be able to achieve the Coulomb blockade effect at room temperature in the future.

In Fig. 3, the effect of the microwave irradiation on I–V_{sd} characteristics is shown. The data were taken at the lowest temperature with the microwave power changing from -40 dBm to -12 dBm, while the lowest one corresponds to the curve without irradiation. We should note that the applied microwave power was kept small such that the temperature would not change during the measurement. Without microwave irradiation, the curve shows the Coulomb gap and structures due to the excited state of the 0D quantum state. As the microwave power is increased, the curve becomes smooth with less structure. The differential conductance, which is calculated from the I–V_{sd} data and shown in the inset of the figure, indicates that the curve become less structured as the microwave power is increased. We have made measurements at different frequencies and at 4.2 K, and observed the similar result. These experimental observations suggest the classical coupling mechanism of high-frequency radiation to the dot. The high-frequency field couples to the source–drain voltage as an ac source in addition to the dc bias voltage. The measured dc current under microwave irradiation is the time-averaged current over the modulating voltage around the fixed dc bias voltage. This may explain the observed I–V_{sd} characteristics under microwave irradiation. At the moment, we do not fully understand the reason of the classical coupling even through hf for 20 GHz ($\sim$1 K) is larger than kT ($\sim$0.25 K) [8–10]. Since special care was not paid to the connection between the coaxial line and the sample gate, the electromagnetic field around the dot could be very much complicated. The microwave signal may couple to various electrodes with macroscopic dimensions, resulting in the effective modulation of V_{sd}.

In summary, we have measured electrical transport of a single bundle of SWNTs at low temperature. The basic characteristics can be understood by that of a single quantum dot separated from the source and drain electrodes with tunneling barriers. The charging energy and the level spacing are larger by an order compared
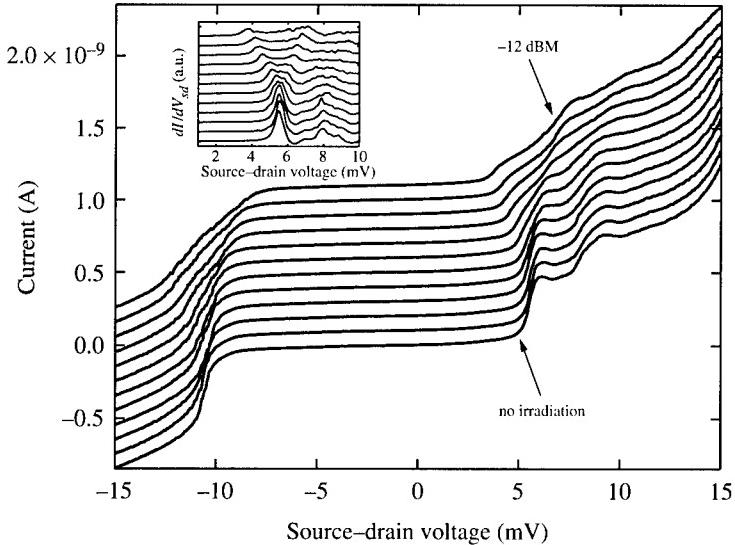
Fig. 3. I–V_{sd} characteristics at $T_{mix} = 25$ mK under 20 GHz irradiation with different microwave powers at a gate voltage of -200 mV. The curves correspond to microwave power levels of no irradiation (bottom), -40, -37, -35, -30, -27, -24.5, -22, -19.5, -17, -14.5 and -12 dBm. Inset: differential conductance calculated from the I–V_{sd} data.

with those for a conventional surface gate GaAs dot. By reducing the separation between the electrodes, the dot could be made smaller and higher-temperature operations could be expected. However, the origin of the tunnel barrier formation has to be investigated in the future study. We also have presented the preliminary result on the microwave irradiation effect on the dc transport of the dot. It is found that the microwave signal effectively modulates the source–drain voltage in a classical manner. Further optimal design of the electrodes is necessary for the efficient coupling of the microwave signal to the dot, which may realize photon assisted tunneling.

References

[1] S. Iijima, Nature **354**, 56 (1991).

[2] M. Bockrath, D. H. Cobden, P. L. McEuen, N. G. Chopra, A. Zettl, A. Thess, and R. E. Smalley, Science **275**, 1922 (1997).

[3] S. J. Tans, M. H. Devoret, H. Dai, A. Thess, R. E. Smalley, L. J. Geerligs, and C. Dekker, Nature **386**, 474 (1997).

[4] M. Bockrath, P. H. Cobolen, J. Lu, A. G. Rinzler, R. E. Smalley, L. Balents, and P. L. McEuen, Nature **397**, 598 (1999).

[5] K. Tsukagoshi, B. W. Alphenaar, and H. Ago, Nature **401**, 572 (1999).

[6] *Proceedings of the NATO Advanced Study Institute on Mesoscopic Electron Transport*, edited by L. L. Sohn, L. P. Kouwenhoven, and G. Schön (Kluwer Academic Publishers, Dordrecht, Netherlands, 1997).

[7] A. T. Johnson, L. P. Kouwenhoven, W. de Jong, N. C. van der Vaart, and G. J. P. M. Harmans, Phys. Rev. Lett. **69**, 1592 (1992).

[8] R. J. Fitzgerald, J. M. Hergenrothen, S. L. Pohlen, and M. Tinkham, Phys. Rev. **B57**, 9893 (1998).

[9] L. P. Kouwenhoven, S. Jauhar, J. Orenstein, P. L. McEuen, Y. Nagamune, J. Motohisa, and H. Sakaki, Phys. Rev. Lett. **73**, 3443 (1994).

[10] T. H. Oosterkamp, L. P. Kouwenhoven, A. E. A. Koolen, N. C. van der Vaart, and C. J. P. M. Harmans, Phys. Rev. Lett. **78**, 1536 (1997).

Interface and contact structures for nanoelectronic devices using assemblies of metallic nanoclusters, conjugated organic molecules and chemically stable semiconductor layers

D. B. JANES, M. BATISTUTA, S. DATTA, M. R. MELLOCH[†]

School of Electrical and Computer Engineering, Purdue University, W. Lafayette, IN 47907, U.S.A.

R. P. ANDRES, J. LIU

School of Chemical Engineering, Purdue University, W. Lafayette, IN 47907, U.S.A.

N.-P. CHEN, T. LEE, R. REIFENBERGER

Department of Physics, Purdue University, W. Lafayette, IN 47907, U.S.A.

E. H. CHEN, J. M. WOODALL

Department of Electrical Engineering, Yale University, New Haven, CT 06520, U.S.A.

Self-assembly ('building') approaches can provide well-controlled structures and assemblies at the nanometer scale, but typically do not provide the specific structures or functionalities required for robust nanoelectronic circuits. One approach to realize high-density nanoelectronic circuits is to combine self-assembly techniques with more conventional semiconductor device and circuit approaches ('chiseling') in order to provide suitable functionality and arbitrary circuit functions. An interesting challenge is to find approaches where these techniques can be combined to realize suitable device structures. This paper describes recent work which combines self-assembly techniques involving metal nanoclusters and conjugated organic molecules with semiconductor interface and device structures to form structures of interest for nanoelectronics. One key requirement for this approach is the availability of a chemically stable semiconductor surface layer, which can provide a low-resistance interface between the metallic nanostructure and the semiconductor device layers following room-temperature, *ex situ* processing. As an illustration of the structures which can be realized, we describe a nanometer-scale ohmic contact to n-type GaAs which utilizes low-temperature-grown GaAs as the chemically stable interface layer. Contact structures have been realized using both isolated (sparse) clusters and using close-packed arrays of clusters on the surface. The low-resistance contacts between the nanoclusters and the semiconductor device layers indicates that relatively low surface barriers and high doping densities have been achieved in these *ex situ* structures. The general conduction model

---

[†]Also at: NSF MRSEC for Technology Enabling Heterostructure Materials, Purdue University, W. Lafayette, IN 47907, U.S.A.

for this contact structure is described in terms of the interface electrical properties and the contributions from the various components are discussed.

The traditional way to fabricate microelectronic devices and circuits involves 'chiseling', i.e. the application of selective deposition or removal of material in lithographically defined areas. Alternatively, it is possible to 'build' well-controlled structures at the nanometer scale using chemical self-assembly techniques. There are a number of self-assembly techniques which have been developed to provide uniform nanometer-scale elements and assemblies of these elements. Of particular interest to this study are structures comprised of metal or semiconductor nanoclusters [1–3]. An interesting challenge is to find approaches where the structures available from 'building' can provide functionality comparable to that realized by 'chiseling' of semiconductor circuits. This paper describes recent work which combines self-assembly techniques involving metal nanoclusters and conjugated organic molecules with semiconductor interface and device structures to form structures of interest for nanoelectronics. We have developed *ex situ* fabrication techniques which are compatible with the self-assembly techniques, as well as suitable low-resistance interface/contact structures at the nanometer scale. We will present a brief summary of the various components and will describe the fabrication, performance and general conduction model for a low-resistance nanoscale ohmic contact structure which has been constructed using this approach.

The general approach described in this paper combines self-assembled nanostructures formed from metallic nanoclusters and conjugated organic molecules with suitable semiconductor heterostructures to realize structures of interest for nanoelectronic applications. The structure of the low-resistance nanoscale ohmic contact which is constructed using this approach is illustrated in Fig. 1. Since the chemical self-assembly techniques are generally not compatible with high-temperature post-processing, and since the successful interfacing of nanometer-scale metallic clusters with the surface requires a well-defined nanometer-scale interface, an essential feature for this hybrid approach is the availability of a chemically stable semiconductor surface layer which does not rapidly oxidize and which remains electrically active during air exposures of at least several minutes. In the current study, surface layers of low-temperature-grown GaAs (LTG:GaAs) [4], i.e. GaAs grown by molecular-beam epitaxy at substrate temperatures in the range of 250–300 °C, are used to provide the requisite surface stability.

The particular structures utilized in this study are based on Au nanoclusters which are approximately 4 nm in diameter. These clusters can be formed into uniform 2D arrays, and the conductance between adjacent clusters can be varied by changing the conjugated organic molecules which are used to link adjacent clusters [1]. The clusters are synthesized in an aerosol reactor [5], and are typically coated with an alkanethiol such as dodecanethiol in order to prevent agglomeration when the clusters are placed into a colloidal suspension. Individual clusters are fcc single crystals in the shape of a truncated octahedron. The well-defined facets of these gold nanocrystals can be observed in high-resolution transmission electron microscope (TEM) images; their influence on the structure of assemblies can also be observed in arrays of the clusters when relatively short linking molecules are used to join adjacent clusters [6]. In contrast to gold clusters which are synthesized in aqueous solution as charged particles, these clusters are charge neutral, which facilitates the formation of closely packed cluster arrays and eliminates offset charge problems when interfaced with device structures. The use of a mono-thiol/unconjugated encapsulant such as dodecanethiol also allows the insertion of conjugated di-thiol linking molecules between adjacent clusters or between a cluster and a semiconductor (or metal) surface, as is accomplished in the work reported here.
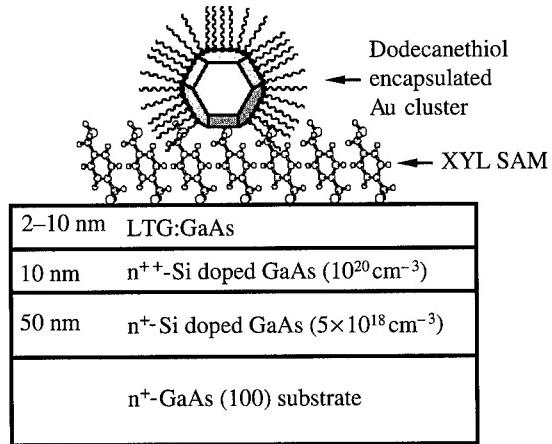
Fig. 1. A schematic diagram of the nanocontact structure utilizing a 4 nm diameter Au cluster, a self-assembled monolayer of xylyl dithiol and a GaAs heterostructure with a chemically stable surface layer.

Conjugated organic molecules that are difunctional such as aryl dithiols can provide mechanical linking/tethering between a cluster and another surface (a cluster or a semiconductor/metal surface) as well as a path for electronic conduction between the respective conductors. The specific molecule used for linking the metal nanoclusters to the semiconductor surface in this study is xylyl dithiol ($HS-CH_2-C_6H_4-CH_2-SH$) which will be referred to as XYL for brevity. This molecule has a length of approximately 1 nm and has thiol (–SH) end groups on each end. These thiol groups can chemically bond to the GaAs surface and to the gold nanoclusters. A self-assembled monolayer (SAM) of XYL is grown on the GaAs surface; given the size of the molecule and the expected packing density, it is believed that a number of XYL molecules (approximately 30) are packed within the area of a facet on a 4 nm diameter Au cluster. The interface between a nanocluster and the GaAs surface therefore consists of a small bundle of XYL molecules.

XYL is one of a class of conjugated organic molecules that have been studied as molecular conductors [7–9]. Such experimental studies typically involve the formation of a self-assembled monolayer (SAM) of the desired molecule on a relatively flat Au surface, with a top contact formed either by metal evaporation, deposition of a metal nanocluster, or by a scanning probe tip (e.g. a scanning tunneling microscope (STM) tip). Theoretical descriptions of the current–voltage relationships have been developed based on the scattering theory of transport. With proper treatment of the Fermi level position in the molecule, of the coupling strength to the contacts and of the capacitive division of the applied voltage, the low-field resistances and current–voltage characteristics for the various experimental reports can be adequately described by this modeling approach [7, 9]. If strong bonding (i.e. strong electronic coupling) to the metallic contacts is realized, vertical transport through a SAM of a short molecule such as XYL should have a specific resistance on the order of 1×10^{-8} Ω cm^2.

Since the self-assembly approaches involving nanoclusters and conjugated organic molecules cannot withstand the process temperatures typically associated with the annealing steps of alloyed contacts or activation of implanted dopants, it is essential that suitable semiconductor device structures are employed to provide low-resistance interfaces between the nanostructure and the doped semiconductor layers. In order to illustrate the need for a suitable surface, consider how ohmic contacts are typically made to n-type GaAs. Since it is difficult to make a contact structure in which the conduction band in n-type GaAs lines up with the Fermi level in the contact metal, tunneling type contacts are typically employed. In principle, it is simple to design a low-resistance tunneling contact: the activated donor density near the surface should be as high as possible and the surface barrier as low as possible, presumably through the use of a relatively low work-function

metal. In practice, these goals are rather difficult to achieve in *ex situ* contacts due to the limited activated donor densities which can be obtained in bulk GaAs and to the mid-gap surface Fermi level pinning which is associated with rapid oxidation of GaAs upon exposure to air. In large-area devices, alloyed contacts such as Au/Ge/Ni are used to overcome these limitations, at the expense of rather deep and nonplanar metal to semiconductor interfaces. Such approaches are not suitable for nanometer-scale ohmic contacts, particularly those involving a nanocluster/semiconductor interface, due to the nonplanar nature of alloyed contacts and to the limited thermal budget dictated by the nature of the self-assembled structures.

Large-area, low-resistance nonalloyed contacts to n-type GaAs layers have been demonstrated using semi-conductor heterostructures comparable to that illustrated in Fig. 1 [10]. It is informative to describe the performance and mechanisms of the large-area contacts in order to illustrate the essential features of the semiconductor layers in the nanocontact structure. In the large-area contact structures, the LTG:GaAs layer thickness is generally between 2 and 5 nm and the contact metal (typically Ti) is deposited by evaporation shortly after an oxide-strip step. The large-area contact is therefore a metal–semiconductor (M–S) structure. The ohmic contact structure [10] employs a surface layer of 'as-grown' LTG:GaAs, in which the $\sim$1–2% *excess* arsenic incorporated during growth is distributed primarily as arsenic antisite defects. The high concentration ($\sim$1.0 $\times$ 10^{20} cm^{-3}) of point defects results in short minority carrier lifetimes and bulk Fermi level pinning [4]. These defects are observed as a band of states located approximately mid-gap in the GaAs [11]. Previous studies indicate that these states prevent the GaAs surface from rapidly oxidizing due to the relatively low concentration of minority carrier holes in the surface layer [12, 13]. As a result, the presence of the gap states can be observed using STM even following brief air exposure of the samples [12].

In the large-area contact studies, the *ex situ*, nonalloyed ohmic contacts employing a LTG:GaAs surface layer and Ti metallization can provide specific contact resistivities (ρ_c) as low as 3 $\times$ 10^{-7} Ω cm^2 [10]. Applications of this contact to shallow device layers and studies of their temperature stability have been reported [14]. Since these contacts do not suffer from the deep interface and spatial nonuniformity of alloyed contacts, they are of interest for nanometer-scale device applications. This type of contact structure and the chemically stable LTG:GaAs surface layer are also compatible with chemical self-assembly techniques, particularly the nanocluster/molecule-based structures employed in this study [1].

The nanocontact structure illustrated in Fig. 1 employs the same basic heterostructure as the large-area nonalloyed contact, but the nature of the interface between the metal and the semiconductor surface are different in the two cases. In particular, the nanocontact requires a molecular tether layer (the XYL) in order to provide suitable mechanical stability. Without this layer, the clusters are too mobile to form a stable contact structure. As will be described later, the XYL monolayer also plays an important role in achieving a low-barrier height and low tunneling resistance through the nanocontact. The nanocontact structure is therefore best described as a metal–'insulator'–semiconductor (M–I–S) structure, in which the XYL represents a very thin, and relatively leaky, insulator. The thickness of the low-temperature-grown GaAs surface layer is 10 nm for the nanocontact structures described in this work. In these studies, a controlled area nanocontact is formed by a single crystal, 4 nm diameter Au nanocluster deposited on the GaAs surface, which has previously been coated with a SAM of XYL. Details of the procedure have been presented elsewhere [15, 16]. The area of the conducting path for the nanocontact (1 $\times$ 10^{-13} cm^2) is defined by the area of a facet on the 4 nm diameter cluster. The current–voltage characteristics of the nanocontact structures have been measured using ultra-high-vacuum (UHV) STM current–voltage spectroscopy in the near-contact regime, i.e. a regime in which the resistance between the STM tip and the cluster is reduced by bringing the tip in close proximity to the cluster [15, 16]. STM tips used in this study have typical end shapes with diameters less than 15 nm, as observed from TEM micrographs. Representative current–voltage curves are shown in Fig. 2A for a nanocontact consisting of an isolated cluster, with curves shown for cases where the tip is over the cluster ('A'), i.e. conduction through the cluster-based nanocontact structure, and where the tip is over a region of XYL coated GaAs, but not over a cluster ('B'). As illustrated in the figure, $I(V)$ curves measured over the cluster indicate a significant enhancement in the relative conduction at low bias voltages as compared
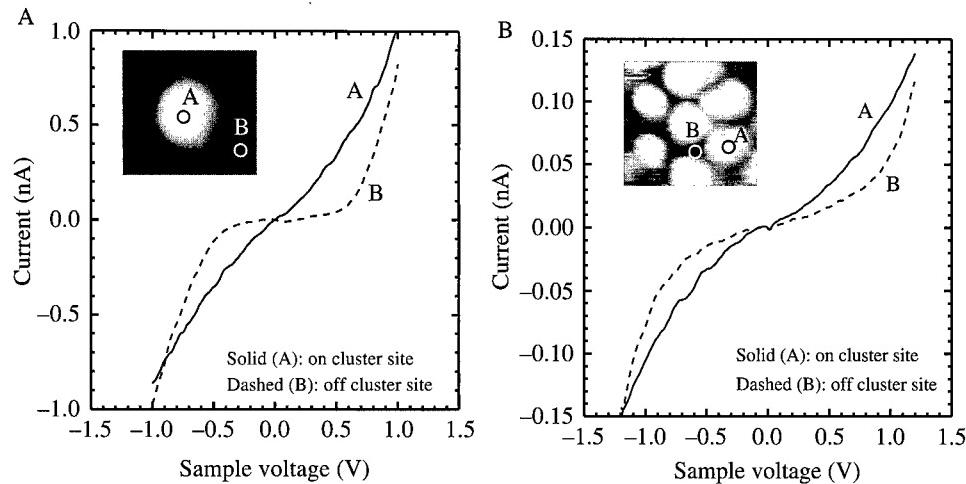
Fig. 2. Measured current–voltage relationships from UHV STM characterization of nanocontact structures with cross-section depicted in Fig. 1. In each figure, curves are shown for I–V relationships both on a cluster and over the XYL-coated surface, with the respective curves corresponding to measurements at the locations indicated on the inset images. A, Representative curves acquired at a set-point of $I_{set} = 0.8$ nA, $V_{set} = -1.0$ V for an isolated nanocluster tethered to the XYL-coated heterostructure, as illustrated in the inset (a 20 × 20 nm STM topographic image of Au cluster, acquired with $I_{set} = 1.0$ nA and $V_{set} = -1.0$ V). B, Representative curves acquired at a set-point of $I_{set} = 0.15$ nA, $V_{set} = -1.2$ V for a structure with a close-packed array of Au nanoclusters on the XYL-coated heterostructure, as illustrated in the inset (a 25 × 25 nm STM topographic image of array of tethered clusters, acquired with $I_{set} = 0.1$ nA and $V_{set} = -1.2$ V).

with the corresponding curves for the measurement over the XYL coated surface (without cluster). The curves shown in Fig. 2 are taken at relatively low set-point currents, corresponding to relatively large tip-to-sample distances. In order to determine the resistance of the nanocontact, i.e. the resistance between the Au nanocluster and the doped layers in the semiconductor structure, it is necessary to reduce the tip-to-cluster distance, and thereby the tip-to-cluster resistance. This has been accomplished by monitoring the current (I) at constant voltage as the tip height (z) is decreased with respect to the height at the low-current set-point. For a structure in which the LTG:GaAs layer is undoped (n-type), the measured $I(z)$ relationship saturates as the tip is brought closer to the cluster. In the saturation regime, the overall resistance should be dominated by the cluster-to-substrate resistance, so the desired nanocontact resistance can be determined from the ratio of the applied voltage to the the saturation current [15]. For this case, the specific contact resistance is approximately 1×10^{-6} Ω cm^2 and the maximum current density is approximately 1×10^6 A cm^{-2} [15]. For a structure in which the LTG:GaAs layer is heavily doped with Be, as is the case in Fig. 2, values of ρ_c of approximately 1×10^{-7} Ω cm^2 and the maximum current density of approximately 1×10^7 A cm^{-2} are obtained. Both of these values are comparable to the values achieved in high-quality large-area ohmic contacts to n-type GaAs [16]. In this case, the measurement system limits the minimum ρ_c which can be resolved, so the contact resistance may be somewhat lower than this value. Thus, the measured specific contact resistances for the best nanocontact structures are somewhat better than those achieved in the large-area contact studies, indicating that high-performance nanocontacts can be formed to GaAs device layers using this nonalloyed ohmic contact approach.

We have also characterized structures in which the clusters are deposited as close-packed arrays on the XYL coated LTG: GaAs surface. A STM topographic image of an array of 5 nm diameter clusters formed on the XYL-coated LTG:GaAs structure is shown in Fig. 3. This image illustrates the good local ordering achieved through self-assembly; the faceted nature of the clusters can also been seen. Representative I–V curves are shown in Fig. 2B, again for the case where the STM tip is over a cluster ('A') and for the case where
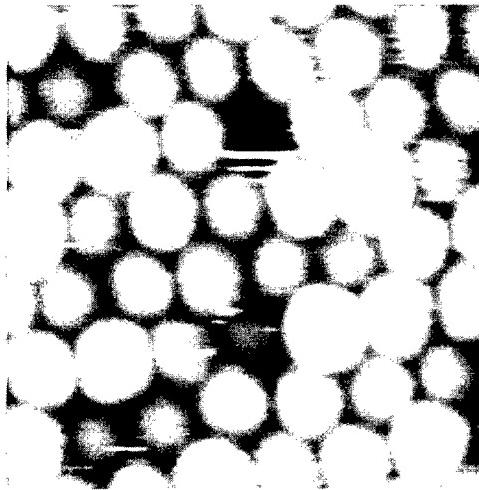
Fig. 3. A 50 × 50 nm UHV STM topographic image of close-packed 2D array of Au nanoclusters tethered to the XYL-coated semiconductor heterostructure with Be-doped LTG:GaAs surface layer, acquired with $V_{set} = -1.2$ V and $I_{set} = 0.1$ nA.

the STM tip is positioned between clusters ('B'). As was observed in the experiment with isolated clusters, an enhancement in low-field conduction is observed when the tip is positioned over a cluster. The 'set-points', i.e. the current/voltage pairs used to establish the relative tip height with respect to the local surface, are approximately the same in the 'on-cluster' and 'off-cluster curves', in order to allow reasonable comparison between the low-field current values. While the nature of the STM technique makes it difficult to ascribe absolute conductance values to the data, the fact that these trends persist to relatively high current levels (when the tip is brought closer to the cluster) indicates that the conduction is enhanced when the conduction is through a cluster. The current–voltage relationships for clusters within an array are comparable to those obtained on isolated clusters. This observation is consistent with the fact that the cluster array used in this study is 'unlinked', i.e. adjacent clusters within this array are separated by the dodecanethiol encapsulant but not linked by conjugated molecules. In this case, the coupling to the semiconductor substrate is much stronger than the intercluster (resistive) coupling. If adjacent clusters within the array were linked with a conductive molecule, it is expected that the intercluster resistance could be made comparable to the cluster-to-substrate resistance. In this case, the conduction path to the substrate would be through a number of clusters. The difference in shapes of the $I(V)$ curves when the STM tip is located on and off a cluster are not as dramatic in the case of the cluster within an array. Since the end size of the tip is 10–15 nm, it is likely that there is some conductivity through adjacent clusters for the 'off-cluster' curve in the array sample (Fig. 2B).

As a starting point for modeling the nanocontact performance, consider a quantitative conductance model for the large-area ohmic contact (M–S structure) which has recently been developed [17]. This analysis calculates the conduction band profile for the semiconductor structure (solution of Poisson equation and Fermi statistics), using parameters for the mid-gap and shallow acceptor states in the LTG:GaAs which are consistent with experimental observations. A calculated profile for a contact structure with a LTG:GaAs layer thickness of 3 nm, a barrier height of 0.5 eV and an activated donor density of 1×10^{20} cm^{-3} is shown in Fig. 4. Also shown in the figure is the band profile for a uniformly doped Schottky barrier (assuming complete ionization) with the same barrier height and comparable depletion depth. Since this profile is reasonably approximated by a parabolic profile, the contact resistance can be calculated by applying expressions for a uniformly doped Schottky contact with equivalent barrier heights and effective depletion depths. This model adequately predicts experimental trends for specific contact resistance versus LTG:GaAs layer thickness and
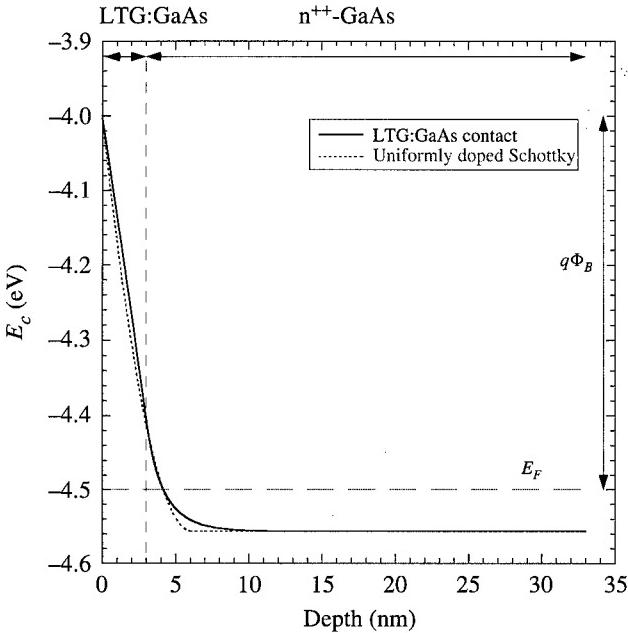
Fig. 4. Calculated conduction band profile for ohmic contact structure (metal/GaAs) with a 3 nm thick LTG:GaAs layer on the surface. The curve is for a structure with a barrier height of 0.5 eV, a doping density in the n^+-GaAs layer of 1×10^{20} cm^{-3} and a temperature of 300 K. Also shown is the parabolic conduction band profile for an equivalent uniformly doped Schottky barrier, assuming complete ionization.

versus temperature. Based on comparisons between the experimental results and the predictions, it appears that the surface barrier heights are well below mid-gap values and that activated donor densities above the bulk amphoteric limit have been achieved in the space charge region [17]. It is typically not possible to achieve either of these effects in *ex situ* contacts to n-type GaAs, due to the surface Fermi level pinning associated with rapid surface oxidation and the bulk amphoteric doping limit in stoichiometric layers. The achievement of both effects in this structure is attributed to the passivating effects of the thin LTG:GaAs layer and to the associated Fermi level control. The ability to achieve an activated donor density within the surface space-charge region which is higher than the bulk amphoteric limit has been explained in terms of the control of Fermi level during growth and the passivation effects of the LTG:GaAs layer [10]. It should be noted that the semiconductor heterostructure is typically exposed to air for prolonged periods before contact processing, so a portion of the LTG:GaAs layer does oxidize during this storage period. Since this oxidation appears to cause mid-gap surface Fermi level pinning even in LTG:GaAs following prolonged air exposure [18], the surface oxide must be stripped shortly before contact metallization. It is believed that the oxide strip restores the surface Fermi level to an unpinned state, with a re-oxidation time constant of hours.

For the nanocontact structure (M–I–S), this conduction model must be modified to incorporate the effects of the layer of XYL between the cluster and the GaAs surface. Qualitatively, this layer can be described as a leaky dielectric, since there are states associated with the molecular levels (HOMO and LUMO states). Generally, one of these two levels will be within 0.5–1 eV of the Fermi level when the molecules are sandwiched between two metallic electrodes [7]. Since the dielectric constant of this layer (expected to be around 2) is considerably lower than that of GaAs, a relatively large portion of the electrostatic potential difference between the metal cluster and the semiconductor bulk is dropped across the XYL layer. This effect reduces the barrier at the surface of the GaAs, again with respect to the GaAs bulk, and makes the barrier relatively more

transparent in comparison with a case without the organic monolayer. This effect is especially important in light of the fact that the workfunction of Au is higher than that of the Ti metallization used in large-area contacts, which would tend to raise the barrier height in the M–I–S with respect to that of the M–S structure. In order to explain the low specific contact resistance in both the large-area contact and in the nanocontact, it is necessary to assume a relatively low density of interface charge at the GaAs surface. Since the surface oxide is chemically stripped from the semiconductor surface immediately before growth of the XYL monolayer, it is believed that the surface Fermi level is also unpinned in the nanocontact structure. Independent observations of the relative stabilities of organic monolayers on undoped and Be-doped layers of LTG:GaAs indicate that the Be-doped layers are more stable. Therefore, it is likely that the nanocontact samples with Be doping in the LTG:GaAs surface layer have a lower interface state density than those in the samples with undoped surface layers. This trend could explain the superior contact resistance properties of the samples with Be-doped surface layers.

In addition to the cluster-based nanocontacts, we have demonstrated other structures which utilize either the chemically stable LTG:GaAs surface layer for device structures or patterned organic monolayers on GaAs, which can provide templates for approaches which combined self-assembly for well-ordered nanometer-scale assemblies and lithographically based techniques for arbitrary global ordering [19, 20]. These structures and approaches represent a toolbox which may eventually be suitable for self-assembling a device structure with functionality comparable to that of a transistor.

A number of devices have been reported in which some feature is in the nanometer scale, including a few recent examples [21–24]. Frequently the overall device dimensions are much larger than this minimum feature size, particularly in structures employing semiconductor channels and source/drain regions. This size discrepancy is largely due to the need for ohmic contact structures which are in the micrometer scale in lateral extent and typically 100 nm or greater in depth. In contrast, an ohmic contact technology which can provide nanometer contact dimensions, both laterally and vertically, could allow the demonstration of the high circuit densities promised by nanometer-scale device concepts.

In conclusion, we have described a nonalloyed contact structure which may be suitable for high-density nanoelectronic device applications and discussed several experiments aimed at developing nanoscale functional devices. In the nanocontact structure, the controlled dimension contact is formed by a gold nanocluster with well-defined crystal facets. Strong mechanical tethering of the cluster to the semiconductor surface is provided by a self-assembled monolayer of a conjugated organic molecule (xylyl dithiol). The low-resistance contact between the nanocluster and the doped semiconductor layer is attributed to the chemical stability of the low-temperature-grown GaAs surface layer, which results in a low interface density, and to the passivation effects and conduction properties of the xylyl dithiol. The ability to form stable, low-resistance interfaces between metallic nanoclusters and semiconductor device layers using *ex situ* processing allows chemical self-assembly techniques to be utilized to form interesting nanoscale semiconductor devices. These demonstrations provide device approaches and fabrication techniques which can be integrated to develop a high-density nanoelectronic device technology with high-throughput fabrication processes.

References

[1] R. P. Andres, J. D. Bielefeld, J. I. Henderson, D. B. Janes, V. R. Kolagunta, C. P. Kubiak, W. Mahoney, and R. G. Osifchin, Science **273**, 1690 (1996).
[2] C. B. Murray, C. R. Kagan, and M. G. Bawendi, Science **270**, 1335 (1995).

[3] S. Facsko, T. Dekorsy, C. Koerdt, C. Trappe, H. Kurz, A. Vogt, and H. L. Hartnagel, Science **285**, 1551 (1999).

[4] M. R. Melloch, J. M. Woodall, E. S. Harmon, N. Otsuka, F. H. Pollak, D. D. Nolte, R. M. Feenstra, and M. A. Lutz, Annu. Rev. Mater. Sci. **25**, 547 (1995).

[5] R. S. Bowles, J. J. Kolstad, J. M. Calo, and R. P. Andres, Surf. Sci. **106**, 117 (1981).

[6] J. Liu, J. Du, and R. P. Andres, unpublished.

[7] W. Tian *et al.*, J. Chem. Phys. **109**, 2874 (1997).

[8] M. Reed, Proc. IEEE **87**, 652 (1999).

[9] S. Datta *et al.*, Phys. Rev. Lett. **79**, 2530 (1997).

[10] M. P. Patkar, T. P. Chin, J. M. Woodall, M. S. Lundstrom, and M. R. Melloch, Appl. Phys. Lett. **66**, 1412 (1995).

[11] R. M. Feenstra, J. M. Woodall, and G. D. Pettit, Phys. Rev. Lett. **71**, 1176 (1993).

[12] S. Hong, D. B. Janes, D. McInturff, R. Reifenberger, and J. M. Woodall, Appl. Phys. Lett. **68**, 2258 (1996).

[13] T. B. Ng, D. B. Janes, D. McInturff, and J. M. Woodall, Appl. Phys. Lett. **69**, 3551 (1996).

[14] H. J. Ueng, V. R. Kolagunta, D. B. Janes, K. J. Webb, D. T. McInturff, and M. R. Melloch, Appl. Phys. Lett. **71**, 2496 (1997).

[15] T. Lee *et al.*, Appl. Phys. Lett. **74**, 2869 (1999).

[16] T. Lee, N.-P. Chen, J. Liu, D. B. Janes, E.-H. Chen, M. R. Melloch, J. M. Woodall, R. P. Andres, and R. Reifenberger, Appl. Phys. Lett. **76**, 212 (2000).

[17] N.-P. Chen, H. J. Ueng, D. B. Janes, K. J. Webb, J. M. Woodall, and M. R. Melloch, to appear in J. Appl. Phys. (2000).

[18] T. Holden *et al.*, Phys. Rev. **B58**, 7795 (1998).

[19] D. B. Janes *et al.*, J. Vac. Sci. Technol. **B17**, 1773 (1999).

[20] J. Liu, T. Lee, B. L. Walsh, R. P. Andres, D. B. Janes, M. R. Melloch, J. M. Woodall, and R. Reifenberger, to appear in Appl. Phys. Lett. (2000).

[21] D. L. Klein, R. Roth, A. K. L. Lim, A. P. Alivisatos, and P. L. McEuen, Nature **389**, 699 (1997).

[22] A. N. Korotkov, *Coulomb Blockade and Digital Single-electron Devices*, edited by J. Jortner and M. Ratner (Blackwell Science Ltd., 1997).

[23] D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. A Kastner, Nature **391**, 156 (1998).

[24] H.-I. Lee, S.-S. Park, D.-I. Park, S.-H. Hahm, J.-H. Lee, and J.-H. Lee, J. Vac. Sci. Technol. **B16**, 762 (1998).

Spin-polarized transport in carbon nanotubes

K. TSUKAGOSHI[†‡], B. W. ALPHENAAR[†]

[†]*Hitachi Cambridge Laboratory, Madingley Road, Cambridge CB3 0HE, U.K.*
[‡]*The Institute of Physical and Chemical Research (RIKEN), Wako, Saitama, 351-0198, Japan*

We present transport measurements of ferromagnetically contacted carbon nanotubes. In both single- and multi-walled nanotube devices, a spin valve effect is observed due to spin-polarized transport. In one single-walled nanotube device, the spin-valve effect is suppressed as the influence of Coulomb charging is observed at around 10 K. To help understand the interplay between the Coulomb charging and the spin-polarized transport we investigated the temperature dependence of the carbon nanotube magnetoresistance.

© 2000 Academic Press

Key words: carbon nanotube, spin-polarized transport.

1. Introduction

The carbon nanotube is a promising candidate as a component in nanoscale molecular electronic devices [1–9]. It acts as a molecular wire, with conducting properties that are far better than a metal wire of similar dimensions. Because the carbon nanotube is a stable, self-assembled structure, its width is uniform, and it is relatively defect-free [1–3]. In addition, metallic nanotubes show extremely high conductivity, and a long mean-free path. Ballistic transport has been observed in nanotubes even at room temperature [4]. The spin scattering length in carbon nanotubes is also expected to be quite long, opening the possibility of creating a carbon nanotube magnetoelectronic device [5, 6]. In this report, we explore this possibility and present measurements of electron spin transport in carbon nanotubes.

2. Sample preparation and experimental details

We disperse nanotubes onto a SiO_2/Si substrate, and map out the position of nanotubes with respect to Pt/Au alignment marks on the substrate using a scanning electron microscope. We use multi-walled carbon nanotubes (MWNTs) or single-walled carbon nanotubes (SWNTs). Contact patterns are defined so as to overlap the mapped nanotube using electron-beam lithography, after which 65 nm of Co is deposited by thermal evaporation at a pressure of 4×10^{-7} Torr. The resulting polycrystalline Co film behaves ferromagnetically and has a room-temperature resistivity of approximately 22 $\mu\Omega$ cm. After lift-off, the ferromagnetic leads are connected to nonferromagnetic bond-pads.
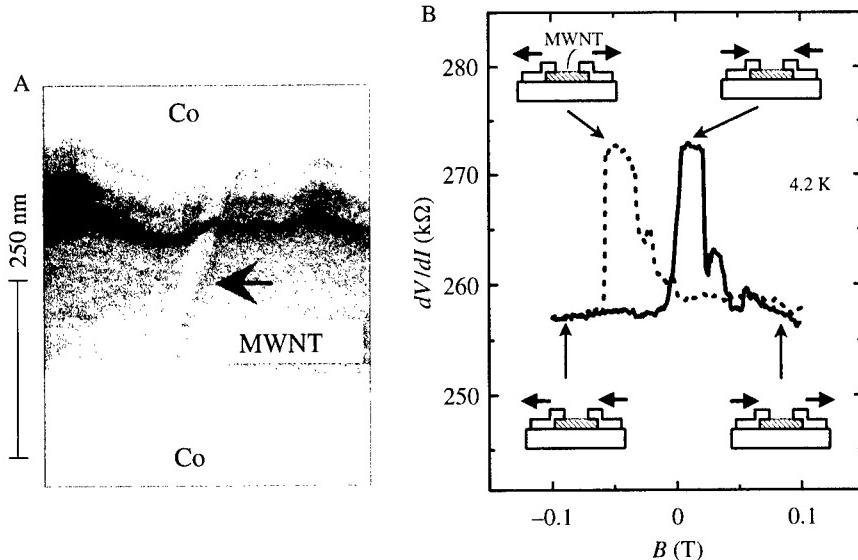
Fig. 1. A, Micrograph of a spin-valve device consisting a single multi-walled carbon nanotube (MWNT) electrically contacted by ferromagnetic Co. The Co contacts lie on top of the MWNT, and the conducting channel is approximately 250 nm in length. B, Two-terminal differential resistance as a function of magnetic field. The magnetic field is directed parallel to the substrate, and the temperature is 4.2 K. The solid (dashed) trace corresponds to the positive (negative) sweep direction. The magnetization direction of the left and right contacts is represented by the direction of the arrows in the figure. The percentage difference $\Delta R / R_0$ between the tunnel resistance in the parallel and the antiparallel states is approximately 6%.

3. Results and discussion

3.1. Multi-walled carbon nanotube

Crude MWNTs synthesized from graphite rods by the arc discharge evaporation method under He atmosphere [2] are used. This ensures that the MWNTs contain no trace of magnetic impurities. The MWNTs are typically 10–50 nm in diameter and a micron or more in length. Figure 1A is an electron micrograph showing the junction region of a completed device.

Magnetoresistance measurements are performed in a 4.2 K bath cryostat with the B-field directed in the plane of the substrate. The two-terminal resistance is measured using an ac lock-in technique with an excitation voltage of 100 μV. Figure 1B shows the two-terminal differential resistance of a Co-contacted nanotube as a function of magnetic field. The field is swept first from −100 mT to 100 mT (solid line) and then back to −100 mT (dashed line). A resistance peak appears as the magnetic field moves through 0 T. There is also a large hysteresis in the peak position (±50 mT) between positive and negative sweep directions, indicating the probable influence of the contact magnetization.

Similar hysteretic magnetoresistance is observed in magnetic tunnel junctions (MTJs), where it has been attributed to spin-polarized electron tunneling [10–15]. The MTJs consist of two ferromagnetic contacts separated by a thin oxide layer. The conduction electrons within the ferromagnetic contacts have a preferred spin direction, which is determined by the local magnetization. This causes the formation of majority and minority spin conduction bands with different densities of states at the Fermi energy. In the absence of spin-scattering, the resistance across the tunnel barrier is dependent on the relative alignment of the magnetization of the two contacts. In the anti-parallel state the majority spin states are out of alignment and the junction resistance is higher than in the parallel state in which the majority spin states are aligned. For the nanotube
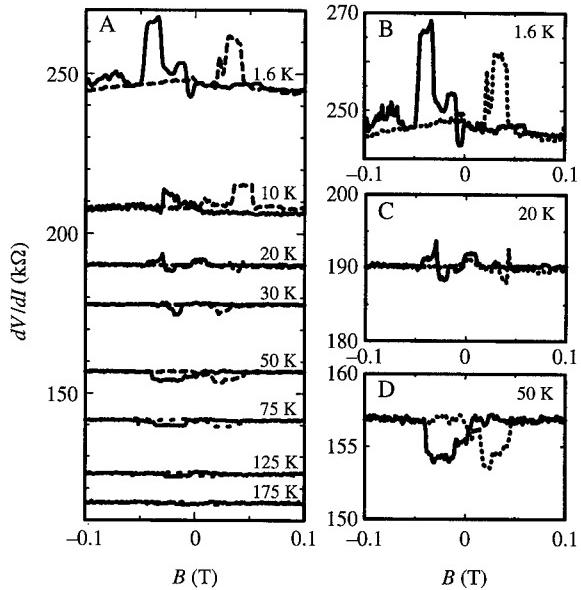
Fig. 2. A, Temperature dependence of the two-terminal differential resistance as a function of magnetic field. B, At low temperatures, the positive spin-valve signal comes out. C, The signal almost disappears at 20 K and D, flips to negative at higher temperature. The spin-valve signal eventually disappears at 175 K.
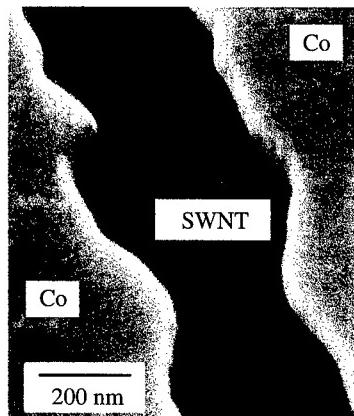
Fig. 3. Micrograph of a single bundle of single-walled carbon nanotubes (SWNTs) electrically contacted by ferromagnetic Co. The Co contacts lie on top of the SWNT, and the conducting channel is approximately 250 nm in length. A side-gate is located at 5 μm apart from the nanotubes.

devices, the contact magnetizations align parallel with the magnetic field at $B = 100$ mT and -100 mT (Fig. 1B). As we sweep B through 0 T, the magnetization polarity switches. The observed peak suggests that the contact magnetizations switch separately and become misaligned as the field is swept. For a MTJ, misalignment occurs because different ferromagnetic contact materials are used, with different coercivities— the magnetizations are misaligned when B lies between the coercive fields of the two contacts. This does not explain the misalignment in the nanotube device. The misalignment may be caused by magnetization fluctuations that occur locally, on the scale of the nanotube diameter (30 nm). The average Co domain size
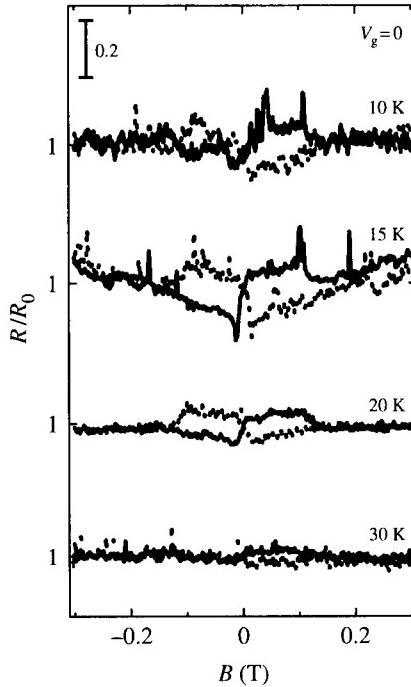
Fig. 4. Two-terminal differential resistance normalized by the zero-field resistance as a function of magnetic field at various temperatures. When the magnetic field is applied parallel to the contact plane, the magnetoresistance has a peak associated with the spin-valve effect between the two ferromagnetic contacts. The resistances are normalized by each zero-field resistance; 2610 kΩ for 10 K, 290 kΩ for 15 K, 175 kΩ for 20 K, and 130 kΩ for 30 K. Solid lines and dotted lines are taken in different field-scan directions. The gate voltage is 0.

(50 nm) [16] is on the order of the width of the nanotube so that the nanotube contacts only a small number of magnetic domains. The coercivity of each domain varies, and depends on its geometry and the local energy conditions.

The sample-to-sample variations observed are probably due to inherent random variations in the surface condition over the small nanotube contact area. Previous experiments on nonmagnetically contacted nanotubes have observed large variations in the contact resistance [7]. Also, in the ferromagnetically contacted samples it is impossible to control the particular domain structure in contact with the nanotube.

The spin-injection picture for the nanotube magnetoresistance requires that a sufficiently small amount of spin scattering occurs both within the nanotube, and at the interfaces between the nanotube and the contacts. Following Julliere's model based on the magnetic tunnel junction [13], the spin-scattering length $l_s \sim$ 130 nm for our best result. Although fairly long, this is probably an underestimation. The spin-polarization near the ferromagnet/nanotube interface will depend on the interface quality.

In the large variations of the domain walls in the polycrystalline Co film, we found a device which shows striking temperature dependence (Fig. 2). As shown in Fig. 2B, the positive magnetoresistance flips to negative at about 20 K while the resistance peak generally becomes smaller as the temperature increases. The polarity-flip was repeatable even after thermal cycles. We suppose that the flip may be attributed to the reconstruction of the domain contributing to the transport in one of the ferromagnetic electrodes. The domain walls pinned at low temperature could be released at higher temperatures to achieve more stable domain walls in the electrodes. Another interpretation uses a model of the two dominant domains in one side of the contacts—one domain has aligned magnetization with the magnetic field and the other is misaligned to the magnetization in the first contact—the two have different temperature dependences. For this model, the
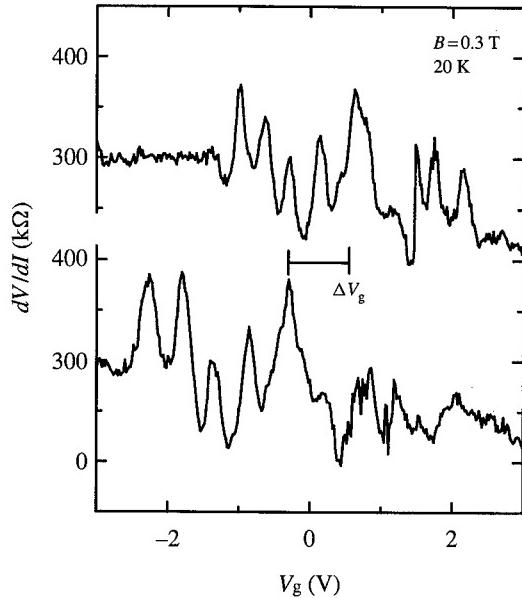
Fig. 5. Gate–voltage characteristics of the ferromagnetically contacted SWNTs with a magnetic field of 0.3 T at 20 K. Upper (lower) trace shows the part of a scan from +4 to −4 V (from −4 to +4 V). The resistance shows oscillations. ΔV_g shows peak shift, which depends on the voltage sweep direction. The value of ΔV_g depends on the applied magnetic field.

domain with misaligned magnetization under the magnetic field survives at higher temperature than the aligned magnetization.

3.2. Single-walled carbon nanotube

For the SWNTs, the transport is more complicated. In the SWNT two-terminal device with a metal contact, single-electron tunneling was observed [8, 9].

We use bundles of SWNTs (less than 10 nm in bundle diameter) synthesized by the pulsed laser vaporization of a metal/carbon target [3]. The typical length of a bundle is 1–20 μm. The polycrystalline Co contacts are formed on the SWNTs in a similar manner to that for the MWNTs. The contact separation is approximately 250 nm (Fig. 3). In the Co-SWNTs device, the spin-valve effect was observed. Figure 4 shows two-terminal differential resistance as a function of magnetic field at various temperatures. In general, the signal of the spin-valve effect becomes more pronounced at lower temperature. In the Co-SWNTs device, the magnetoresistance change, at first, becomes larger nearer to 15 K. At lower temperature, however, the signal was suppressed, and the conductance oscillations, which might be caused by the Coulomb blockade effect, become pronounced in the gate–voltage characteristics. This may indicate that spin polarity of the conduction current is randomized during tunneling through the Coulomb island.

In order to see the co-existing state with the spin coherence and the Coulomb blockade effect, we measured the gate–voltage characteristics at 20 K under the magnetic field of 0.3 T. The resistance oscillates in the gate–voltage characteristics (Fig. 5). The observed peak is hysteretic in gate–voltage scan; the shift of ΔV_g is about 1 V. The reproducible hysteresis with the same oscillations period is observed while we keep the temperature and the applied magnetic field constant. The ΔV_g is largest near the zero-field, and saturated under the higher magnetic fields. Thus, this might be not the conventional current shift due to electron traps near the single electron transistor. We found similar hysteretic gate–voltage characteristics in other device.

The result is appears to be peculiar to carbon SWNTs. Interpretation of this result, however, will require a further investigation.

References

[1] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes* (Imperial College Press, Singapore, 1998).

[2] T. W. Ebbesen and P. M. Ajayan, Nature **358**, 220 (1992).

[3] A. Hess *et al.*, Science **273**, 483 (1996).

[4] S. Frank *et al.*, Science **280**, 1744 (1998).

[5] K. Tsukagoshi, B. W. Alphenaar, and H. Ago, Nature **401**, 572 (1999).

[6] B. W. Alphenaar, K. Tsukagoshi, and H. Ago, Physica E **6**, 848 (2000).

[7] A. Bachtold *et al.*, Nature **397**, 673 (1999); C. Shöenberger *et al.*, Appl. Phys. **A69**, 283 (1999).

[8] S. J. Tans *et al.*, Nature **394**, 761 (1998).

[9] M. Bockrath *et al.*, Science **275**, 1922 (1997).

[10] P. M. Tedrow and R. Meservey, Phys. Rev. Lett. **26**, 192 (1971).

[11] A. G. Aronov, Pis'ma Zh. Eksp. Teor. Fiz. **24**, 37 (1976) [JETP Lett. **24**, 32 (1976)].

[12] M. Johnson and R. H. Silsbee, Phys. Rev. Lett. **55**, 1790 (1985).

[13] M. Julliere, Phys. Lett. **54A**, 225 (1975).

[14] J. S. Moodera *et al.*, Phys. Rev. Lett. **74**, 3273 (1995).

[15] T. Miyazaki and N. Tezuka, J. Magn. Magn. Mater. **139**, L231 (1995).

[16] U. Rüdiger *et al.*, Phys. Rev. **B59**, 11914 (1999).

Interface interactions and the photoluminescence from asbestos-templated InP quantum wires

E. A. ZHUKOV

M.V. Lomonosov Moscow State University, Physics Department, 119899 Moscow, Russia

H. M. YATES, M. E. PEMBLE

Department of Chemistry, University of Salford, Salford M5 4WT, U.K.

C. M. SOTOMAYOR TORRES

Institute of Materials Science and Department of Electrical Engineering, University of Wuppertal, 42097 Wuppertal, Germany

S. G. ROMANOV

Institute of Materials Science and Department of Electrical Engineering, University of Wuppertal, 42097 Wuppertal, Germany
Ioffe Physical Technical Institute, Politekhnicheskaya 26, 194021 St Petersburg, Russia

The photoluminescence of InP quantum wires embedded in the dielectric template has been studied by site-selective pump-probe spectroscopy. The enhancement of the emission in the spectral window around the pump frequency and the long-lasting memory of pumping have been observed and explained on the basis of changes in energy states of template defects under intensive pumping. The effect of wire–template interface upon the appearance of the photoluminescence spectrum is discussed.

© 2000 Academic Press

Key words: quantum wires, templated nanostructures, interfaces, pump-probe photoluminescence.

Embedding semiconductors in nanochannels of crystalline dielectric templates results in three-dimensional ensembles of uniform quantum wires (QWRs) [1]. The advantages of this approach are a high density of nanostructures, crystal-like ordering and a deep potential well. Combining the size quantization and dielectric confinement, such materials provide an increase of the binding energy and the oscillator strength of excitons over that achieved with all-semiconductor heterostructures. Generally, the binding energy increases with the dielectric constant mismatch of the wire and the template due to the stronger Coulomb interaction between electrons and holes in the QWR surrounded by the insulator [2–4]. Exciton binding energies in the range of 100–200 meV have already been observed in GaAs and InP QWRs stabilized in the channels of chrysotile asbestos [5–7].

The behaviour of QWRs embedded in an insulator appears complicated because any electronic process involving the excitation and relaxation within the electronic system of the guest QWR is mirrored by the distortion in the electronic balance of the host medium. The template–semiconductor interface represents a continuous plane of defects, which embraces the QWR and separates materials of different properties. We

have addressed the problem of interface interaction earlier by comparing the absorption and photolumines-cence (PL) spectra of InP QWRs embedded in asbestos [8] and other channel templates [9]. In the present study we use site-selective PL spectroscopy [10] in its pump-probe mode to characterize the interface effect upon the radiative relaxation of electron–hole pairs.

Chrysotile asbestos (CA) is a naturally occurring serpentine mineral with $Mg_3Si_2O_5(OH)_4$ composi-tion [11]. The spiral roll of double silica–magnesium planes forms fibres possessing an empty central channel of 5 nm diameter. Fibres are arranged hexagonally and stuck together in bundles, where the channel diameter variation is within 10% of the mean value. Fibres in studied samples are about 1 cm long and 40 nm wide. Asbestos represents a unique template owing to its optical transparency, thermal stability and mechanical rigidity.

InP was synthesized in asbestos channels by sequential low-temperature metallorganic chemical vapor deposition (MOCVD). The synthesis starts with the absorption of trimethylindium (TMIn) in the channels and is followed by phosphine in a flow of nitrogen [12]. The adsorption centres are smoothly distributed over the channel surface providing the homogeneous loading.

PL spectra were excited by a probe beam at 2.54 eV with 1 mW power concentrated in a spot of 200 μm diameter on the sample cooled down to 2 K. The intensity of the probe beam was kept low to avoid the saturation of the PL spectrum. The pumping was performed in the same spot with frequencies from 1.823 to 2.066 eV using a dye laser with the linewidth of about 0.1 meV and the intensity of 200 mW. Illumination was performed alternately by the probe and pump beams with the 25 Hz frequency, moreover the entrance slit of monochromator was closed during the pumping period.

Figure 1 shows PL spectra of InP–asbestos and bare asbestos excited by the probe beam. Both spectra show a band at 2.28 eV, which is assigned to oxygen defects of the silica skeleton of asbestos [8]. Another weak PL band of defects appears as a shoulder at 1.89 eV, it is shown by dash line subtracted from the background. The 1.9 eV band of InP–asbestos has been interpreted as the radiative recombination of excitons in InP QWRs [7]. PL spectra in the probe and pump-probe regime are shown in Fig. 2. Pumping in resonance with the InP exciton band gives rise to a PL intensity enhancement by factor of 2 at the high-energy side of the PL band accompanied by a 1.4 times suppression at the low-energy side. A blue shift of 13 meV of the PL peak occurs with respect to the pump frequency. PL spectrum returns slowly to that of the unexposed sample after finishing the pumping (Fig. 2). PL response of InP–asbestos to the probe excitation under strong pumping with different wavelengths is shown in Fig. 3A. Spectral features similar to those in Fig. 2 are always present, although their appearance depends on the pump frequency. With the aim of further analysis, the differential spectra obtained by subtracting the PL spectrum excited by the probe beam from those obtained in the pump-probe regime are plotted in Fig. 3B.

The electron energy distribution in InP QWRs is close to, and partly overlaps with, that of asbestos (as shown in Fig. 1). In addition, the QWRs and the asbestos matrix are in contact in real space. Apparently, one can expect a potential barrier between QWRs and asbestos, which does not allow free carrier exchange between them. However, tunneling between QWRs and template becomes possible if: (i) the electrochemical potential of QWRs is levelled with that of asbestos and (ii) the level of the electrochemical potential of asbestos overlaps the defect band [8].

The absorption of InP–asbestos substantially exceeds that of bare asbestos in the spectral range of inter-est [8]. Consequently, the probe beam excites more efficiently electron–hole pairs in InP than at the oxygen defect sites. Once electrons and holes are generated in QWRs they form excitons aided by confinement in the deep potential well [4]. The radiative recombination of excitons in QWRs produces the broad PL peak at 1.9 eV. The broadening has several sources including the dispersion of QWR diameters and the interaction between electron–hole pairs with the template. The latter is facilitated by oxygen vacancies at the surface of channels, which are effectively positively charged. Thus, QWRs are located in an irregular potential land-scape, which in turn spreads the energy distribution of excitons. The valleys of this landscape trap electrons in QWRs, which results in a reduced probability of radiative recombination.
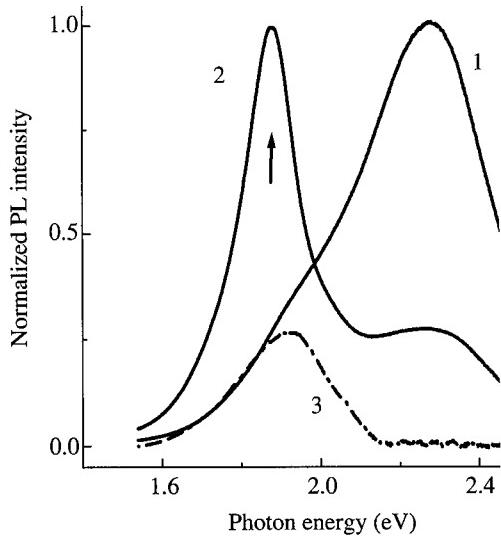
Fig. 1. PL spectra of (1) bare and (2) InP infilled asbestos under 2.54 eV probe excitation. The 1.88 eV PL band of bare asbestos (3) after subtracting the Gaussian approximation to the peak centered at 2.28 eV from curve (1).
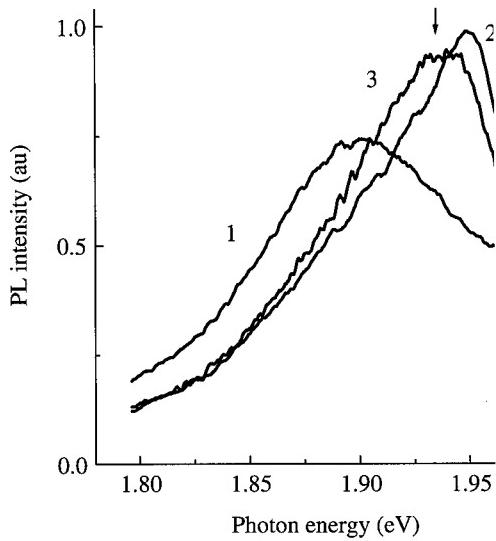
Fig. 2. PL spectra of InP–asbestos excited by the 2.54 eV probe (1), by 1.935 eV pump and 2.54 eV probe (2) and by the probe beam after 40 min since pumping (3). The arrow marks the pump excitation energy.

Irradiation of InP–asbestos with the pump beam in resonance with the exciton band results in a high density of electron–hole pairs in QWRs. Some of them tunnel from QWRs to template defects and saturates defects sites at the interface. This effect is equivalent to transferring a negative charge on these defects. The remaining electrons relax radiatively within a few nanoseconds in a QWR after finishing the pump pulse [7]. In contrast, all excitations in the template are frozen by the potential barrier. This assumption correlates with the observation of long-living changes of PL spectra of the pumped sample, moreover the pump memory is erasable by exposing the sample to light of frequency below the absorption edge of InP
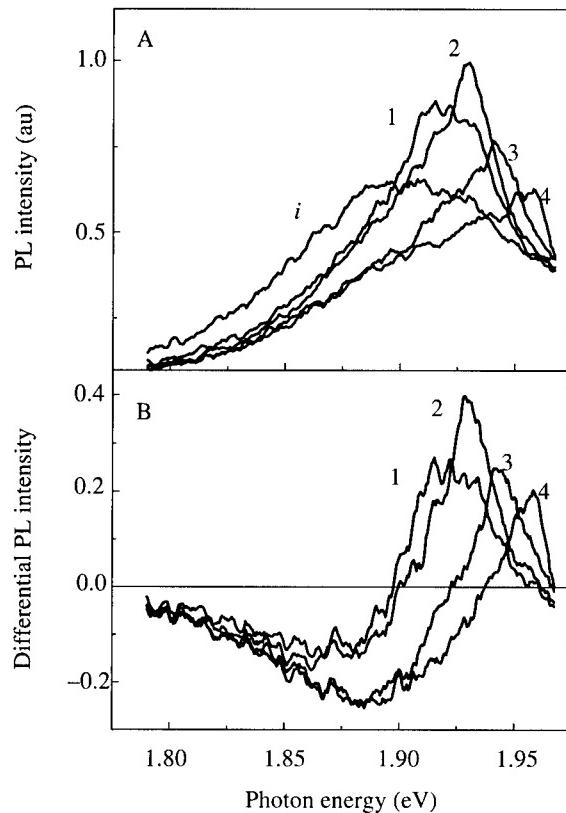
Fig. 3. A, PL spectra in the probe regime (curve *i*) and in the pump-probe regime at pump frequencies of (1) 1.906, (2) 1.925, (3) 1.937 and (4) 1.953 eV for the same intensity of pumping. B, Differential PL spectra obtained by subtracting spectrum (*i*) from pump-probe spectra of corresponding number in the upper panel.

QWRs or by heating the sample. Both ways can activate the return of electrons over the interface barrier. Obviously, saturated traps do not interfere with the relaxation process of photoexcitations created under the subsequent probe beam. This cancellation of one particular nonradiative relaxation channel manifests itself as an increase of the PL intensity under the pump-probe beam excitation. The enhancement of the PL intensity is shown as a peak in the differential spectrum in Fig. 4A. In Fig. 4D it is seen that the area under this peak decreases with pump frequency, probably reflecting the inhomogeneous width of the exciton emission peak.

The width of the enhanced PL peak of the pumped sample greatly exceeds the spectral width of the pump beam because the saturation of traps is mediated by QWRs. The pump radiation generates a high density of electron–hole pairs in QWRs, which interact with each other and spread their energy distribution before tunneling through the interface barrier. Therefore, the probability of radiative recombination for subsequently probe-induced excitons increases within a broad transparency window around the pump frequency. Remarkably, the half-width of the PL enhancement peak measured in two different ways (Fig. 4C) is nearly independent of the pumping frequency.

The decrease of the PL intensity on the low-energy side of the 1.9 eV PL band can be accounted for by the reduced contribution from the template PL to the overall PL signal. The photodarkening effect in the PL from oxygen defects is known to occur in porous silica after intense pumping over the whole spectral range
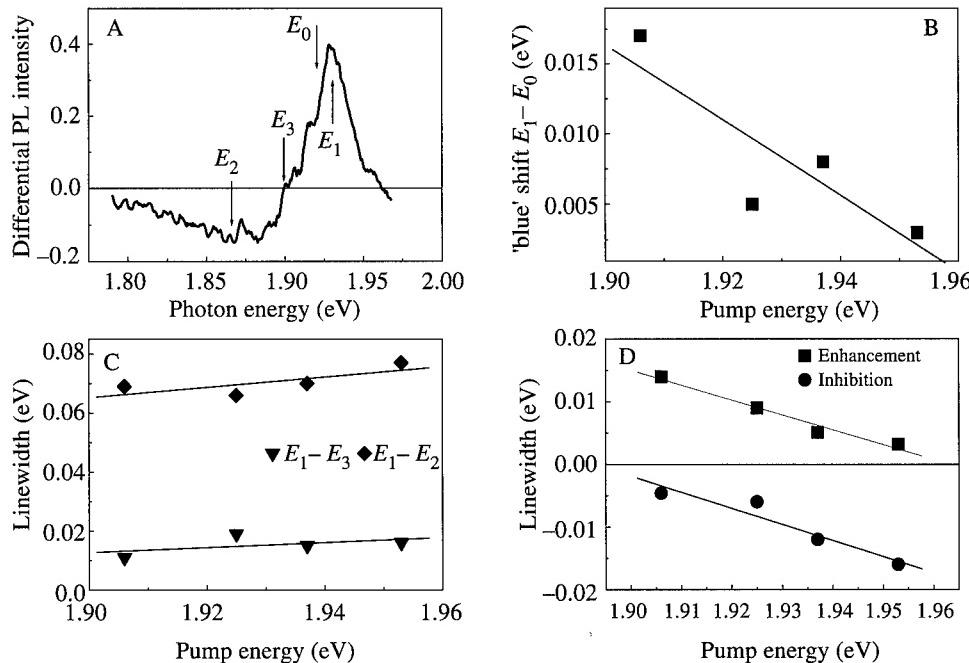
Fig. 4. A, Assignment of characteristic energies in differential PL spectrum: E_0–pump energy, E_1–PL enhancement maximum, E_2–PL inhibition maximum, E_3–energy separation of the enhancement- and inhibition-related bands. B, 'Blue' shift between the PL maximum and the pump frequency. C, Spectral width of the enhancement peak of differential spectra. D, Areas under enhancement and inhibition peaks of differential spectra (from data in Fig. 3B).

below the pumping frequency. This effect is visualized as a dip in differential spectra (Fig. 3B). In agreement with this mechanism, the higher the pumping frequency, the larger the area of this dip (Fig. 4D).

The radial electrical field, which is induced by charges trapped at the interface of QWRs, is proportional to the number of defects and their charge state. Since the number of charged defects is proportional to the area of the PL enhancement peak and assuming that the captured charge density is higher than the density of defects with electron deficit, we may deduce that the magnitude of the electrical field in the pumped volume of the sample exceeds that of the unpumped volume. Due to this field the exciton binding energy in QWRs decreases (Stark effect) and its transition frequency increases, i.e. PL maximum appears 'blue' shifted with respect to the pump frequency (Fig. 4B). Consistently, this shift changes in the same manner as the enhancement area changes (Fig. 4D).

To summarize, two components of the InP–asbestos contribute to the emission, moreover, the QWR–template interface governs their balance in the total spectrum. The strong energy-selective pumping at a frequency resonant with the exciton PL band was used to separate the contributions to the total PL signal and to reveal the actual width of the exciton emission peak. This separation appears possible due to the difference in the timescale of the relaxation processes in QWRs and the template. The main lesson from this study is that the complication of the energy relaxation in the nanocomposite is the cost of the increase of the binding energy of excitons in structurally confined nanoparticles.

References

[1] S. G. Romanov and C. M. Sotomayor Torres, Three-dimensional lattices of nanostructures—template approach, in *Handbook of Nanostructured Materials and Technology*, edited by H. S. Nalwa (Academic Press, 1999), Vol. 4, Chap. 4, pp. 231–323.

[2] V. S. Babichenko, L. V. Keldysh, and A. P. Silin, Fiz. Tverd. Tela **22**, 1238 (1980) [Sov. Phys. Solid State **22**, 723 (1980)]

[3] L. V. Keldysh, Phys. Status Solidi (a) **164**, 3 (1997).

[4] E. A. Muljarov and S. G. Tikhodeev, Zh. Eksp. Teor. Fiz. **111**, 274 (1997) [JETP **84**, 151 (1997)]

[5] V. S. Dneprovskii, E. A. Zhukov, E. A. Muljarov, and S. G. Tikhodeev, Zh. Eksp. Teor. Fiz. **114**, 700 (1998) [JETP **87**, 382 (1998)]

[6] V. Dneprovskii, N. Gushina, O. Pavlov, V. Poborchii, I. Salamatina, and E. Zhukov, Phys. Lett. **A204**, 59 (1995).

[7] E. A. Muljarov, E. A. Zhukov, V. S. Dneprovskii, and Y. Masumoto, to appear in Physical Review **B61**(23), (2000).

[8] S. G. Romanov, C. M. Sotomayor Torres, H. M. Yates, M. E. Pemble, V. Y. Butko, and V. V. Tretiakov, J. Appl. Phys. **82**, 380 (1997).

[9] S. G. Romanov, H. M. Yates, M. E. Pemble, J. R. Agger, M. V. Anderson, C. M. Sotomayor Torres, V. Y. Butko, and Y. A. Kumzerov, Physics of the Solid State **39**, 641 (1997).

[10] Y. Masumoto, J. Lumin. **70**, 386 (1996).

[11] K. Yada, Acta Crystallogr. **23**, 704 (1967).

[12] H. M. Yates, W. R. Flavell, M. E. Pemble, N. P. Johnson, S. G. Romanov, and C. M. Sotomayor Torres, J. Cryst. Growth **170**, 611 (1997).

Large-scale quantum mechanical simulations of carbon nanowires

Madhu Menon[†], Ernst Richter

Department of Physics and Astronomy, University of Kentucky, Lexington, KY 40506-0055, U.S.A.
Center for Computational Sciences, University of Kentucky, Lexington, KY 40506-0045, U.S.A.

Padma Raghavan, Keita Teranishi

Department of Computer Science, University of Tennessee, Knoxville, TN 37996-1301, U.S.A.

The stability of quasi-one-dimensional structures of carbon is investigated using a generalized tight-binding molecular-dynamics scheme. Large-scale simulations are made possible by the parallel implementation of the diagonalization routines. Our results show that these structures can be stable provided that their geometries consist of a core of four-fold coordinated atoms, surrounded by a three-fold coordinated outer surface accommodating one of the most stable reconstructions of bulk diamond structure.

© 2000 Academic Press

Key words: nanotubes, molecular dynamics, clusters, eigenvalues.

The potential applications of carbon nanotubes have aroused much speculation [1–5]. Single-wall carbon nanotubes (SWNT) are now abundantly being produced in experiments. While much excitement has been associated with carbon nanotubes, potential useful applications of other quasi-one-dimensional (QOD) solid forms of carbon cannot be overlooked. Recently, an alternative approach to the synthesis of nanoscale structures based on nanotubes has resulted in the production of nanorods (nanowires). These rods have been found to share the properties of the bulk materials (such as magnetism and superconductivity), suggesting that they might allow the investigation of the effects of confinement and reduced dimensionality on such solid-state properties [6]. Some of the nanorods exhibit unique morphologies, such as helical nanorods [6]. The precise structures of these nanorods are not known at present.

In this work we propose a novel, physically motivated approach supported by a quantum molecular dynamics method to the QOD structures of carbon. We base our approach on the premise that robust nontubular structures of carbon can be made stable by having a four-fold coordinated core surrounded by an outer surface of atoms with three-fold coordination. The models proposed for the QOD geometries for C can be broadly categorized into two distinct classes with various degrees of surface to bulk ratio of atoms. All the structures proposed have been fully optimized without any symmetry constraints using a quantum molecular dynamics method.

The theoretical method used in the present work is the quantum mechanical generalized tight-binding molecular dynamics (GTBMD) scheme of Menon and Subbaswamy [7] that allows for full relaxation of

---

[†]E-mail: super250@pop.uky.edu
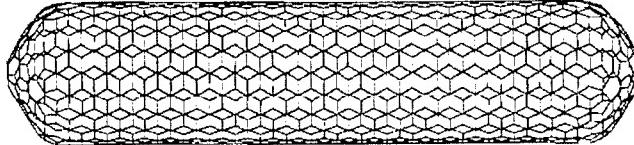
Fig. 1. The GTBMD relaxed (10,10) capsule. Parallel implementation of the GTBMD code was used in relaxing this 1000-atom cluster.

covalent systems with no symmetry constraints. This method has been found to be very reliable in obtaining very good agreement with experimental and local density approximation (LDA) results for the structural and vibrational properties of fullerenes and nanotubes [7]. For details of the formalism we refer the readers to Ref. [7] which also contains applications to the structural and vibrational properties of nanotubes. Briefly, in the GTBMD scheme a system is completely characterized by a parametrized Hamiltonian (H) and an overlap matrix (S). A generalized eigenvalue equation $H\psi = \varepsilon_n S\psi$ is solved to determine the eigenvalues ε_n and eigenvectors ψ of the system. These are then used in the determination of the forces for performing molecular dynamics simulations. The solution of the eigenvalue problem is the computationally most demanding part of the GTBMD method. For a realistic simulations of any system, however, one needs to model the system with a large number of atoms, resulting in a large eigenvalue problem. It is not feasible to solve the associated generalized eigenvalue problem on typical workstations; a parallel solution is essential. It should be noted that both matrices H and S are real symmetric and sparse with sparsity increasing rapidly with the system size. Also, since we are interested mainly in total energies and forces, only half of the eigenvalues and eigenvectors need be evaluated.

The generalized eigenvalue computation is intrinsically expensive; for $N \times N$ matrices, the storage requirements grow as N^2 while the number of operations grow as N^3. Consider a 1061-atom simulation which yields a generalized eigenvalue problem of dimension 4244. On a high-performance workstation using eigenvalue codes from LAPACK [8] the computation time is well over an hour for a single timestep, making the overall simulation infeasible. The simulations reported here were made possible by using a cluster of workstations and dense parallel eigensolvers from ScaLAPACK [9]. These codes are MPI [10] based and are portable across a variety of parallel platforms. The codes use explicit block-cyclic distribution of the dense matrices and cache-efficient matrix operation kernels to reduce overall time. For well-conditioned matrices, the speed-ups are near ideal for small numbers of processors.

We first apply the parallel GTBMD formalism to relax a nanotube 'capsule' obtained by capping the ends of a (10,10) nanotube. The relaxed geometry is shown in Fig. 1 and contains 1000 atoms. Under optimal conditions each timestep of the molecular dynamics simulation took about 5.5 min.

We next consider the QOD structures of carbon. All geometries considered in this work were carefully relaxed to structures that are true local minima of the total energy. The vibrational frequencies were then computed for these relaxed structures within the GTBMD scheme [7] as a further check for stability. None of the structures proposed here had any imaginary frequencies, indicating them to be true local minima of the total energy. Henceforth in this work the term 'locally stable' would indicate the absence of any imaginary frequencies.

The building blocks for class 1 QOD structures are multiply connected 'superatom' clusters with D_{nh} symmetry with a small hollow region in the middle. Figure 2A shows a representative with D_{6h} symmetry. This superatom cluster contains 84 atoms and can be visualized as consisting of six circularly connected closed units of 14 atoms each with faces composed of pentagons and hexagons, with 60 of these atoms having three-fold coordination. The remaining atoms in the unit are all four-fold coordinated. All outer surface atoms participate in forming geometric configurations that resemble very closely the surface of bulk diamond. The bonding of the atoms at the top and bottom, along the symmetry axis, to their three neighbors resemble the unreconstructed (111) surface of diamond. The atoms on the surface farthest from the symmetry axis form
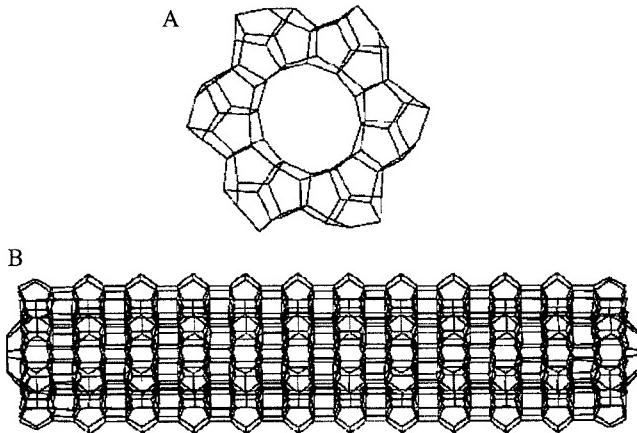
Fig. 2. A, The 'superatom' cluster containing 84 atoms (D_{6h} symmetry). Surface reconstruction results in the formation of symmetric tilted dimers. B, A section of QOD (class 1) wire obtained by stacking the superatom units on top of each other and performing GTBMD relaxation. The ends of the segment show (111) features.

dimers as a result of the 2×1 reconstruction; the unreconstructed surface resembling the (100) surface of diamond. Note that the buckled dimers are all tilted in the same direction (symmetric dimer configuration). The dimer bond length is 1.59 Å. Molecular dynamics relaxations starting with an untilted configuration resulted in the formation of asymmetric tilted dimers. The symmetric tilted dimer configuration shown in Fig. 2A is found to be isoenergetic with the asymmetric case. As seen in the figure, the surface dimers are all aligned along the circumference.

The QOD structures are obtained by stacking these clusters along the symmetry axis. Figure 2B shows a large QOD cluster consisting of 1020 atoms obtained by stacking these units. The stacking results in four-fold coordination for all the (111) type atoms as they form bonds with similar atoms on neighboring unit cells. It is worth noting that (111) surfaces have the lowest surface energy among all other surfaces of diamond. The surface dimers, however, remain three-fold coordinated and do not form bonds with atoms in the neighboring unit cells. The wire-like geometry thus obtained is found to be locally stable. Other class 1 QOD structures of D_{nh} symmetries with $n \neq 6$ can be constructed using the prescription give above.

The class 1 structures can also be used to derive another class of QOD structures with similar structural properties (four-fold coordinated atoms in the interior and three-fold on the outside). In one case (class 2), we consider stable structures with smaller surface-to-bulk ratios. The superatom unit for these classes of QOD structures, consisting of 90 atoms, is obtained from class 1 unit structure by adding a ring of six C atoms around the symmetry axis in the center (Fig. 3A). The GTBMD-relaxed QOD structure contains surface dimers each with a bond length of 1.52 Å and aligned along the circumference. The wire-like structures generated by stacking these clusters are also locally stable. Structure similar to this has also been independently reported by other groups, although no relaxation was performed [11].

In Table 1 we given a summary of relative binding energies. Also, in the same table, the values of the energy difference between the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) are given. As is evident from the binding energies in the table, the structure belonging to class 2 is more stable than that belonging to class 1. The relatively large difference in the HOMO–LUMO gaps between the two classes may be used to help distinguish the two structures experimentally through photo-emission measurements. Interestingly, similar wire structures for Si were also found to be stable in our earlier work [12]. The ordering of energies, however, between classes 1 and 2 were reversed, illustrating the critical differences in bonding in carbon and silicon.

A

B
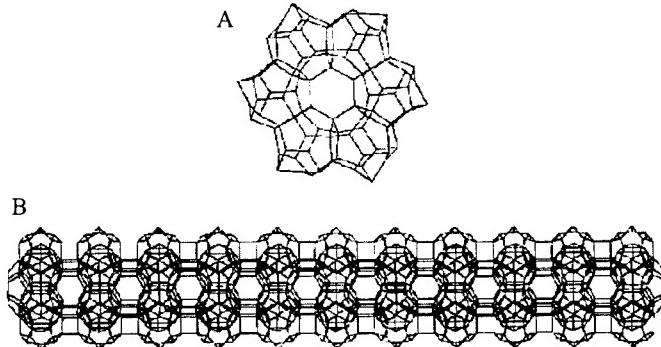
Fig. 3. A, 'Superatom' cluster unit for QOD structures belonging to (class 2) B, and a section of QCD (class 2) wire obtained by stacking the superatom on top of each other and performing GTBMD relaxation.

Table 1: Relative energies of the carbon nanowire clusters obtained using the present scheme. All energies are relative to the (10,10) capsule shown Fig. 1. The values of the energy difference between the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) are also given.

Structure	Energy (eV/atom)	HOMO–LUMO gap (eV)
(10,10) capsule (Fig. 1)	0.00	0.00
Class 1 wire (Fig. 2)	1.08	1.30
Class 2 wire (Fig. 3)	0.62	1.61

The relatively large bond lengths for the three-fold coordinated atoms belonging to classes 1 and 3 may seem counter-intuitive; since in crystalline diamond the surface atoms tend to have shorter bond lengths when compared with the core atoms. A careful examination, however, reveals this 'anomalous' increase in the bond length for surface dimers to be due to the strain caused by the alignment of dimers along the circumference. Also, since the only surface reconstruction for these QOD structures are the 2 × 1 type, we expect the surfaces to be at least as reactive as those of the bulk 2 × 1 surface. The passivation of the surface dangling-bond orbitals through reconstruction is believed to lead to low reactivity.

The presence of large interstitial regions in all these nanostructures may make doping with group III or V elements relatively easy. This may prove to be an important advantage in device fabrication.

In summary, we have proposed novel QOD structures of carbon whose surfaces closely resemble one of the most stable reconstructions of the crystalline diamond with a core of bulk-like four-fold coordinated atoms.

References

[1] R. S. Ruoff, Nature **372**, 731 (1994).
[2] C. Guerret-Piecourt, Y. Le Bouar, A. Loiseau, and H. Pascard, Nature **372**, 761 (1994).
[3] S. C. Tsang, Y. K. Chen, P. J. F. Harris, and M. L. H. Green, Nature **372**, 159 (1994).
[4] P. M. Ajayan and S. Iijima, Nature **361**, 333 (1993).

[5] S. Seraphin, D. Zhou, J. Jiao, J. C. Whithers, and R. Loutfy, Nature **362**, 503 (1993).

[6] H. Dai, E. W. Wong, Y. Z. Lu, S. Fan, and C. M. Lieber, Nature **375**, 769 (1995).

[7] M. Menon, E. Richter, and K. R. Subbaswamy, J. Chem. Phys. **104**, 5875 (1996).

[8] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. J. Dongarra, J. DuCroz, A. Greebaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen, *LAPACK Users' Guide* (SIAM Publications, Philadelphia, PA, 1995).

[9] L. S. Blackford, J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, and R. C. Whaley, *ScaLAPACK Users' Guide* (SIAM Publications, Philadelphia, PA, 1997).

[10] W. Gropp, E. Lusk, and A Skjellum, *Using MPI: Portable Parallel Programming with the Message Passing Interface* (MIT Press, Cambridge, MA, 1994).

[11] E. G. Halpern, V. V. Pinyaskin, and I. V. Stankevich, Russ. Chem. Bull. **45**, 511 (1996).

[12] M. Menon and E. Richter, Phys. Rev. Lett. **83**, 792 (1999).

Implications of atomic-level manipulation on the Si(100) surface: From enhanced CMOS reliability to molecular nanoelectronics

M. C. HERSAM, J. LEE, N. P. GUISINGER, J. W. LYDING

Department of Electrical and Computer Engineering and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A.

The ultra-high vacuum scanning tunneling microscope (UHVSTM) has been used to induce desorption of H from the Si(100)–2X1:H surface with atomic-level precision. The study of the desorption mechanism led to the discovery of a substantial isotope effect between H and D, which has recently been employed to minimize hot electron degradation at the Si/SiO_2 interface in conventional complementary metal-oxide-semiconductor (CMOS) circuits. This paper will reveal secondary ion mass spectroscopy (SIMS) data that show a direct correlation between D incorporation at this interface and transistor lifetime. D incorporation can be enhanced via high-pressure processing, which has led to lifetime improvements in excess of 700× for Samsung's latest 0.18 μm, 1.5 V CMOS technology. In addition to enhancing current integrated circuits, UHVSTM-induced hydrogen desorption has aided the development of nanoelectronics on the molecular-size scale. Feedback-controlled lithography (FCL) has refined the desorption process to the point where templates of individual dangling bonds can be generated in arbitrary geometries. The chemical contrast between dangling bonds and H-passivated Si is then utilized to isolate individual copper phthalocyanine (CuPc) and C_{60} molecules on the Si(100) surface. Following isolation, STM spectroscopy has characterized the mechanical and electrical properties of these molecules with *intra*-molecular precision.

© 2000 Academic Press

Key words: CMOS hot electron reliability, nanoelectronics, STM spectroscopy.

1. Introduction

Nanometer-scale lithography on the Si(100)–2X1:H surface was initially demonstrated by Lyding *et al.* in 1994 [1]. The technique utilized the ultra-high vacuum scanning tunneling microscope (UHVSTM) as a localized electron beam to break Si–H bonds. The desorption mechanism was studied extensively, leading to the understanding of two desorption pathways [2]. For high-energy electrons (>6 eV), direct electronic excitation to the antibonding state prompted the acceleration of H away from the Si surface. On the other hand, low-energy electrons (<3.5 eV) also induced desorption by exciting Si–H bonds through increasingly energetic vibrational states. This model predicts that hydrogen's heavier isotope, deuterium, should be much more difficult to desorb from the Si(100) surface [3]. Experimentally, deuterium is about 50 times and 10^{12} times more difficult to desorb than hydrogen in the direct electronic excitation regime and the vibrational heating regime, respectively [4]. This discovery motivated the replacement of hydrogen with deuterium at

the Si/SiO_2 interface in conventional complementary metal-oxide-semiconductor (CMOS) circuits to minimize hot electron degradation. Initial results on deuterium annealed n-channel metal-oxide-semiconductor (NMOS) transistors showed a lifetime improvement by a factor of 10–50 [5]. This paper updates the latest efforts to improve CMOS reliability with deuterium. Secondary ion mass spectroscopy (SIMS) measurements coupled with NMOS transistor electrical characterization show a correspondence between increased deuterium incorporation at the Si/SiO_2 interface and enhanced transistor lifetime. By increasing the deuterium pressure, the deuterium concentration at the interface can be increased for a fixed annealing temperature. Consequently, equivalent transistor reliability can be obtained with a reduced thermal budget. These high pressure annealing techniques have been applied to Samsung's most recent CMOS technology. For these short-channel (0.18 μm) and low operating voltage (1.5 V) devices, the isotope effect is substantially enhanced. Rather than being scaled away, deuterium reduction of hot electron damage appears more relevant in future generations of CMOS integrated circuits.

Besides impacting modern electronics, STM-induced desorption of hydrogen from silicon is being utilized for futuristic research on molecular nanoelectronics. Since molecular electronics is still in its infancy [6], reliable isolation and characterization of molecules on silicon remains challenging. This paper demonstrates single-atom control of the hydrogen desorption process with a technique called feedback-controlled lithography (FCL). The strong chemical reactivity of the dangling bonds compared with the hydrogen-passivated surface [7] allows for directed self-assembly of deposited molecules onto atomically precise patterns created with FCL. In this study, two species of organic molecules are studied on the Si(100) surface. Two binding configurations of copper phthalocyanine (CuPc) have been observed, and STM spectroscopy allows for the spatial distribution of charge transfer from the silicon substrate to the central copper atom to be detected. Under perturbative scanning conditions, CuPc and adjacent hydrogen atoms can be exchanged along dimer rows. Furthermore, when the CuPc is reacted with ammonia, the resulting molecule weakly interacts with the surface via an outer benzene ring. In this case, the molecule possesses rotational freedom and appears to be spinning. C_{60} has also been isolated following FCL patterning. In this case, STM spectroscopy reveals the lowest unoccupied molecular orbital (LUMO) at an energy of $\sim$1.6 eV above the Fermi level. In addition, intra-molecular spectroscopic structure is observed within the C_{60} molecule as shifts in the LUMO peak of the electronic local density of states (LDOS).

2. Improved CMOS reliability via high-pressure deuterium annealing

Using a homemade high-pressure wafer annealing furnace, fully processed 8″ CMOS wafers can be subjected to deuterium pressures up to 15 atm and temperatures in excess of 450 °C. This furnace has been used to study the effect of increased deuterium pressure on the incorporation of deuterium at the SiO_2/Si interface. The initial experiments were performed on a structure consisting of 600 nm of deposited SiO_2 on a Si(100) substrate. The results of three different deuterium anneals at 450 °C are depicted in Fig. 1. The SIMS data clearly show a direct correlation between deuterium pressure and the resulting deuterium concentration at the interface. In fact, a 6 atm deuterium anneal for 20 min leads to more than an order of magnitude greater deuterium concentration at the interface than a 2 atm anneal for 1 h. Since increased deuterium incorporation at the interface has been directly correlated to improved hot carrier degradation performance [8, 9], high-pressure annealing is expected to increase transistor lifetime. Figure 2 tests this hypothesis by displaying lifetime improvements for two different CMOS technologies following high-pressure deuterium anneals. In both cases, the deuterium results are compared with a 10% H_2, 90% N_2 anneal at 1 atm and 400 °C.

Figure 2A considers Samsung CMOS transistors (0.35 μm channel length, 3.3 V operating voltage) that have been annealed in 100% deuterium at 450 °C following fabrication with four metal layers of interconnections. In this case, the threshold voltage (V_{th}), transconductance (G_m), and source-drain current (I_{ds}) were monitored as a function of time while the NMOS transistors were electrically stressed. The magnitude of improvement relative to the hydrogen passivated sample is recorded in Fig. 2A. For all cases, the deuterium
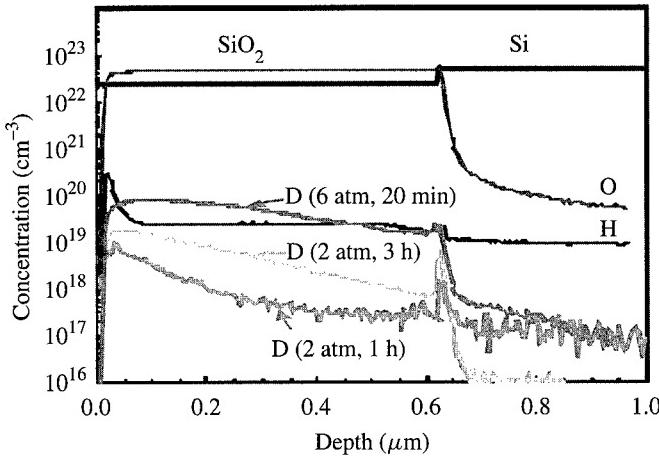
Fig. 1. SIMS results for a 600 nm SiO_2 on Si structure following three different deuterium anneals.

anneal resulted in improved transistor performance under electrical stress. In particular, the 6 atm deuterium anneal for 1 h led to approximately the same degree of improvement as the 2 atm deuterium anneal for 3 h. It should also be noted that the most aggressive annealing conditions (6 atm, 3 h, 450 °C) resulted in a 90× improvement of the threshold voltage, which exceeds the best results from a previous study [10]. In [10], the devices had fewer diffusion barriers and were annealed at 1 atm for 5 h at 450 °C to achieve an 80× improvement. These results show that increased deuterium pressure can reduce the required annealing time for equivalent reliability performance.

In addition to reducing the annealing time, increased pressure can also reduce the annealing temperature. This pressure–temperature relationship is shown in Fig. 2B. In this experiment, the latest Samsung CMOS technology (0.18 μm channel length, 1.5 V operating voltage) was annealed in 100% deuterium after the metal 4 fabrication step. The magnitude of improvement was determined from NMOS transistor lifetime measurements. Clearly, a 6 atm anneal at 400 °C resulted in better reliability performance than a 1 atm anneal at 450 °C for the equivalent amount of time. The enhanced deuterium incorporation rate under high-pressure conditions allows the thermal budget to be reduced. This result may prove to be technologically significant since a reduced thermal budget will enable advanced materials (e.g. low k dielectrics) to be incorporated into future device structures. Furthermore, it should be noted that the maximum lifetime improvement for these transistors was greater than 700×. Compared with the 3.3 V transistors of Fig. 2A, the isotope effect appears greatly enhanced at reduced operating voltages. This observation is consistent with STM desorption experiments where the isotope effect was found to be dramatically increased at low electron energies where the desorption pathway is dominated by excitation via vibrational heating.

3. Electrical and mechanical properties of individual CuPc molecules

As previously discussed, STM-induced desorption of hydrogen has also impacted the emerging field of molecular electronics. By monitoring the STM feedback signal (equivalent to the tip position perpendicular to the surface) during patterning, the desorption of hydrogen can be detected as an abrupt change of 1.5 Å as shown in Fig. 3. Consequently, when patterning under threshold conditions, the termination of the patterning conditions immediately after detecting a desorption event can result in the depassivation of a single silicon atom. This technique has been termed feedback-controlled lithography (FCL) and can create atomically precise arrays of dangling bonds (see Fig. 3). It should be noted that FCL can inherently adapt to subtle

(A) Samsung CMOS (0.35 μm, 3.3 V, Metal 4)

Magnitude of improvement

(B) Samsung CMOS (0.18 μm, 1.5 V, Metal 4)

Magnitude of improvement
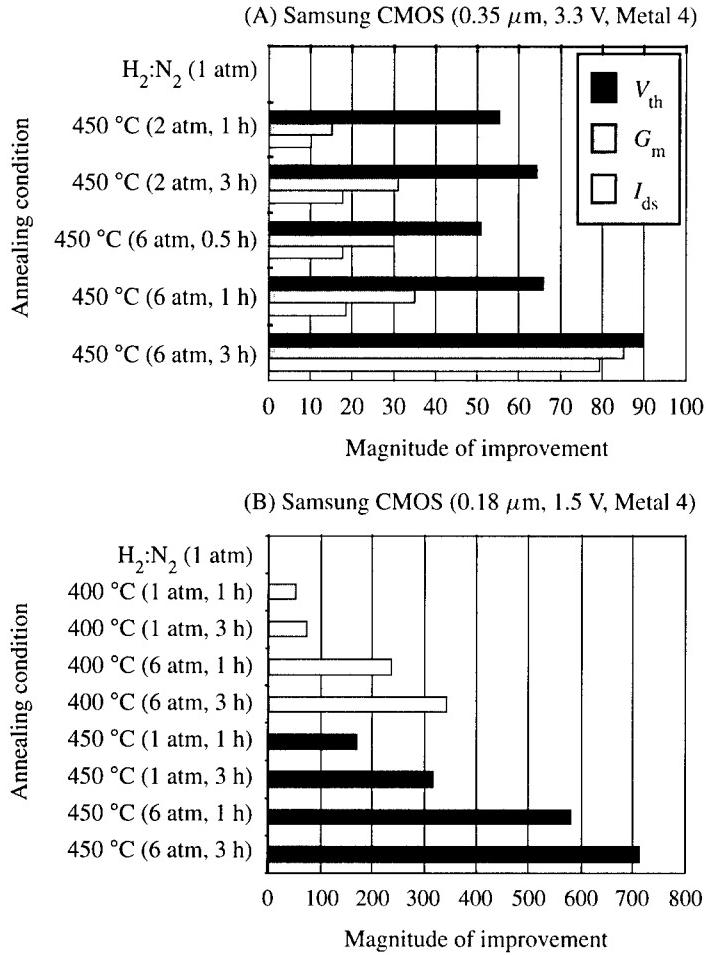
Fig. 2. A, Comparison of three different reliability parameters to a H-passivated device following different deuterium annealing conditions. B, Comparison of transistor lifetime to a H-passivated device following different deuterium annealing conditions.

changes in the tip structure that affect the desorption yield since the desorption time does not need to be known *a priori*.

In Fig. 4A, a V-shaped pattern of dangling bonds was generated using FCL. This patterned surface was then exposed to a sublimed flux of copper phthalocyanine (CuPc). Following a low-dose deposition, individual CuPc molecules were observed at several of the pre-patterned binding sites, as shown in the insets of Fig. 4A. Under empty states imaging conditions, the CuPc molecule appears as a depression with the central copper atom appearing as a protrusion. This behavior has been observed previously [11, 12] and is consistent with the structure of CuPc shown as an inset to Fig. 4B. The other inset to Fig. 4B shows a map of the electronic LDOS obtained at an empty states energy of 2 eV after CuPc deposition. This tunneling conductance map shows a ring of reduced LDOS surrounding each CuPc molecule. This dark ring can be explained as a spatial measure of the charge transfer required from the substrate to reduce the oxidation state of the central copper atom upon reaction with a silicon dangling bond. Figure 4B also illustrates that the STM tip can manipulate individual CuPc molecules. The white-circled dangling bond site of Fig. 4A has been transferred down to the end of the dimer row following CuPc deposition and subsequent STM scanning at -2 V sample bias
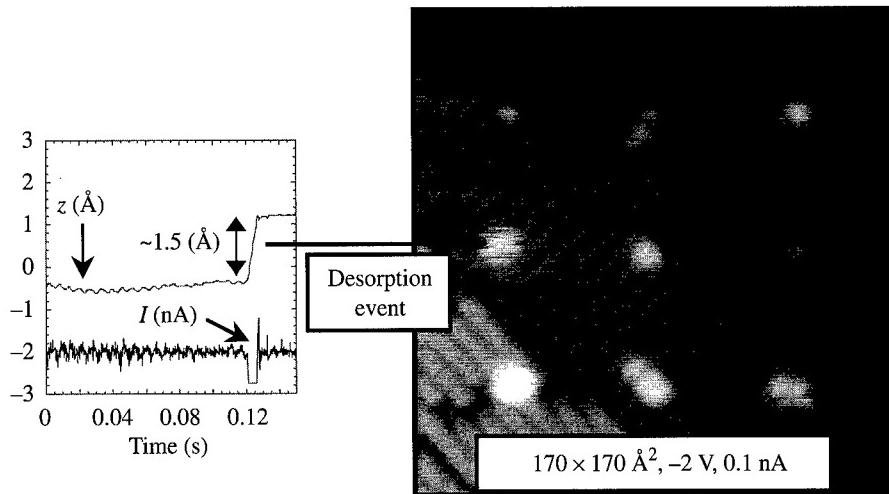
Fig. 3. FCL is used to generate a 3 × 3 array of dangling bonds on the Si(100)–2X1:H surface. The desorption event is detected as a 1.5 Å discontinuity in the tip position (z) after which the patterning conditions of 3 V, 2 nA are terminated.
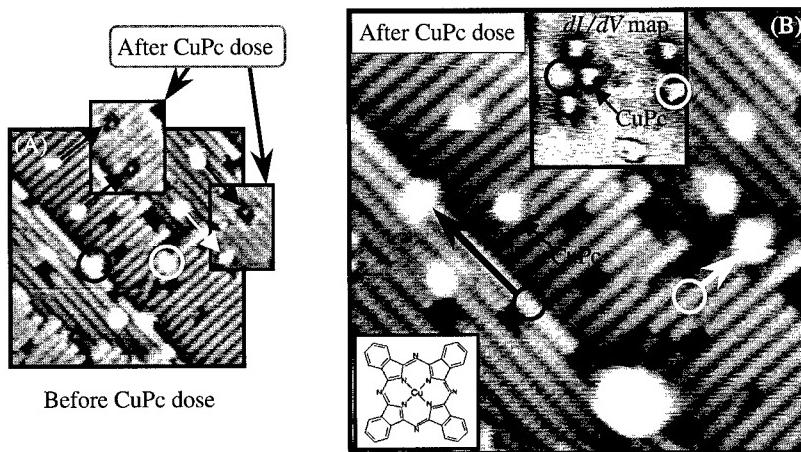
Fig. 4. A, A V-shaped pattern of dangling bonds is generated with FCL. The image size is 170 × 170 Å². The insets show individual molecules bound to the surface following exposure to CuPc. The white arrow indicates that some of the sites remain unreacted following this sub-monolayer dose. B, A filled states image of the surface following CuPc exposure illustrates STM tip induced motion of two of the binding sites. The inset dI/dV map was taken at an empty states energy of 2 eV and shows a ring of reduced LDOS surrounding the CuPc molecules.

and 0.1 nA tunneling current. It should be noted that the previous position of the dangling bond now appears to be hydrogen passivated. Consequently, under these imaging conditions, the STM can occasionally induce the exchange of H atoms and CuPc molecules along dimer rows. A similar behavior has been observed for the black circled dangling bond site of Fig. 4. In this case, however, the tunneling conductance map after manipulation does not possess a ring of reduced LDOS. Apparently, this site no longer possesses a CuPc molecule. This observation implies that the STM can also induce desorption of CuPc from the surface. By scanning under less perturbative conditions (e.g. reduced tunneling current), unintentional manipulation of CuPc by the STM tip can be avoided.
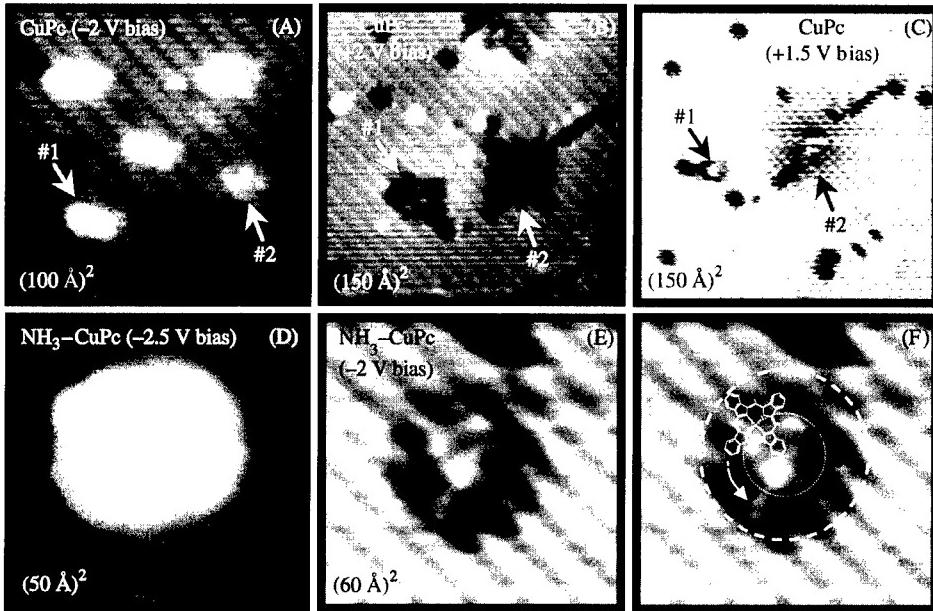
Fig. 5. A–C, A series of STM images taken after a FCL pattern had been exposed to CuPc. Molecule #1 appears to be stationary, whereas molecule #2 is spinning on the surface. D,E, Two STM images of an ammonia-reacted CuPc molecule after isolation on the Si(100)–2X1:H surface. F, A scaled drawing of CuPc has been overlaid on the image from part E, of this figure. This depiction illustrates the 32 Å circle that is swept out by the molecule when it rotates via its outer benzene ring.

Recently, the rotation of a single molecule has been observed on the Cu(100) surface with STM [13]. When CuPc bonds to a silicon dangling bond via its outer benzene ring, similar rotational behavior is detected. In Fig. 5A–C, two CuPc molecules have been isolated using FCL. Figure 5B shows that molecule #1 possesses fourfold symmetry and that the surface appears somewhat brighter surrounding this molecule. Furthermore, in Fig. 5C, molecule #1 possesses the expected 16 Å diameter for CuPc. The fourfold symmetry and 16 Å diameter of molecule #1 suggest that it is held stationary on the surface by a rigid bond with the surface. The apparent brightness of the surface surrounding this molecule is indicative of the charge transfer that was previously discussed for CuPc when it is strongly interacting with a silicon dangling bond via its central copper atom. On the other hand, molecule #2 possess radial symmetry in Fig. 5B and C. Also, the size of molecule #2 in Fig. 5C remains much larger than the expected dimensions of CuPc. A radially symmetric feature with a diameter roughly double the actual size of CuPc suggests that the molecule is rotating via its outer benzene ring at a rate faster than the time resolution of the STM ($\sim$30 kHz). Evidently, the relatively weak interaction between the silicon dangling bond and the benzene ring implies that the molecule has rotational freedom at room temperature. In addition, the surface does not appear brighter surrounding molecule #2 in Fig. 5B, so no charge transfer is occurring in this case as expected.

To further investigate the rotation of CuPc, a derivative of CuPc was synthesized by exposing the original CuPc source material to 120 L of ammonia *in situ*. The ammonia reacts with the central copper atom with the NH_3 group pointing out of the plane of the molecule. Since the NH_3-reacted CuPc molecule already has five ligands attached to the center copper atom, bonding of this molecule to the surface via the center copper atom would require a Jann–Teller distortion to minimize energy for this system. Since the CuPc and surface geometries are confined, this Jann–Teller distortion is not energetically favorable. Consequently, it is expected that this molecule will bond to surface via its outer benzene ring. Figure 5D–F shows that the NH_3-reacted CuPc molecule possesses radial symmetry and an apparent diameter of 32 Å. The inner ring
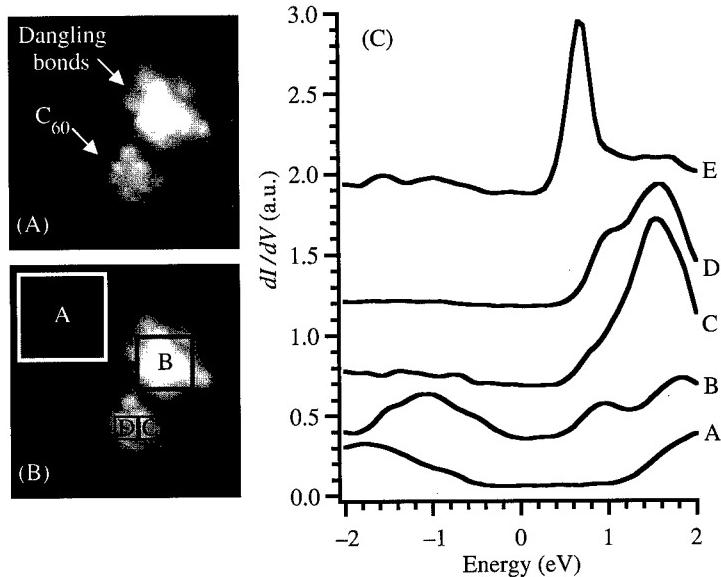
Fig. 6. A, A filled states (70 Å)^2 STM image of an isolated C_{60} molecule next to a group of silicon dangling bonds. B,C, Full STM spectra were taken at every pixel in this image. All of the pixels within each box of part B, were averaged to produce the curves shown in part C. Curves A–E have been offset vertically for clarity.

within the molecule of Fig. 5E can be attributed to the NH_3 group sticking out of the plane of the CuPc. The apparent rotation of this molecule on the surface is consistent with the expectation that it would react with the surface via an outer benzene ring.

4. Intra-molecular spectroscopic structure of C_{60}

To further explore the flexibility of FCL for single-molecule studies, C_{60} has been isolated on the Si(100) surface using this technique. Figure 6A shows an STM topographic image of a C_{60} molecule next to a group of silicon dangling bonds. The intra-molecular structure observed within the C_{60} molecule can be attributed to the spatial distribution of the molecular orbitals at this filled states energy of -2 eV [14]. In this image, full STM spectroscopy was performed at every pixel in the image for energies between ± 2 eV. For these spectroscopic measurements, the tunneling conductance was measured using a lock-in amplifier when the sample bias was modulated with a 20 kHz, 120 mV$_{rms}$ signal.

The five curves of Fig. 6C were obtained by averaging the STM spectra for all of the pixels within each labeled box of Fig. 6B. Box A is located over the hydrogen passivated Si(100) surface. The corresponding dI/dV curve shows a relatively featureless bandgap with the Fermi energy positioned near the valence band edge. Since the silicon used in this experiment was p-type with a resistivity of 0.1 Ω-cm, this Fermi level position is expected. Box B is located over the cluster of silicon dangling bonds. Consequently, curve B shows a reduced bandgap and additional spectroscopic structure beyond the band edges. The most obvious peaks at ± 1 eV are consistent with calculated LDOS curves for the clean Si(100) surface [15]. The agreement of curves A and B with theoretical expectations lends credence to the quality of this spectroscopic data set. As a result, new information can reliably be extracted from the dI/dV curves taken over the C_{60} molecule. Curve C was taken over a region of the C_{60} molecule that is most representative of the average behavior for the entire molecule. The most prominent feature of this curve is a strong peak at about 1.6 eV. Presumably, this peak represents the location of the LUMO with respect to the Fermi energy. Curve D was obtained over the binding

point between the C_{60} molecule and the silicon surface. Therefore, this curve shows the LUMO feature of the C_{60} molecule and the 1 eV peak representative of the underlying silicon dangling bond. Curve E shows that the LUMO peak is shifted within the C_{60} molecule. Although the origin of this shift is not yet understood, this data demonstrates the experimental capability to observe intra-molecular spectroscopic structure.

5. Conclusion

STM-induced hydrogen and deuterium desorption from the Si(100) surface has inspired improved reliability in modern CMOS integrated circuits and led to initial advances in the emerging field of molecular nanoelectronics. High-pressure deuterium annealing of CMOS devices has led to improvements of the incorporation rate of deuterium at the Si/SiO_2 interface. The increased deuterium concentration has been measured with SIMS and correlates directly to reduced hot carrier degradation in NMOS transistors. The enhanced deuterium incorporation rate under high-pressure annealing conditions implies that the annealing time or temperature can be reduced without giving up improved transistor lifetime performance. Such results are potentially significant for future transistor structures where the thermal budget may need to be reduced during fabrication. Finally, the results for Samsung's latest CMOS technology show that the isotope effect is greater in transistors that operate at reduced operating voltages.

An important aspect of future molecular electronic circuits will be the merger of single-molecule switching and storage elements with conventional silicon device technology. To this end, this paper illustrates a technique for isolating and characterizing individual molecules on the Si(100) surface. FCL was introduced and used to create atomically precise templates for CuPc and C_{60} molecules. When CuPc reacts with silicon dangling bonds via its central copper atom, charge transfer is detected from the surrounding silicon surface in tunneling conductance maps. On the other hand, molecular rotation is observed when the molecule binds via an outer benzene ring. On C_{60}, full STM spectra were obtained that revealed the LUMO at an energy of $\sim$1.6 eV above the Fermi level. Furthermore, intra-molecular spectroscopic structure was observed within the C_{60} molecule. With single molecule isolation and characterization demonstrated, nanoelectronic devices can now be explored by interfacing molecules with macroscopic-to-nanoscale electrical contacts [16].

References

[1] J. W. Lyding, T.-C. Shen, J. S. Hubacek, J. R. Tucker, and G. C. Abeln, Appl. Phys. Lett. **64**, 2010 (1994).
[2] T.-C. Shen, C. Wang, G. C. Abeln, J. R. Tucker, J. W. Lyding, Ph. Avouris, and R. E. Walkup, Science **268**, 1590 (1995).
[3] Ph. Avouris, R. E. Walkup, A. R. Rossi, T.-C. Shen, G. C. Abeln, J. R. Tucker, and J. W. Lyding, Chem. Phys. Lett. **257**, 148 (1996).
[4] E. T. Foley, A. F. Kam, J. W. Lyding, and Ph. Avouris, Phys. Rev. Lett. **80**, 1336 (1998).
[5] J. W. Lyding, K. Hess, and I. C. Kizilyalli, Appl. Phys. Lett. **68**, 2526 (1996).
[6] M. A. Reed, Proc. IEEE **87**, 652 (1999).

[7] G. C. Abeln, M. C. Hersam, D. S. Thompson, S.-T. Hwang, H. Choi, J. S. Moore, and J. W. Lyding, J. Vac. Sci. Technol. **B16**, 3874 (1998).

[8] J. Lee, S. Aur, R. Eklund, K. Hess, and J. W. Lyding, J. Vac. Sci. Technol. **A16**, 1762 (1998).

[9] J. Lee, Y. Epstein, A. Berti, J. Huber, K. Hess, and J. W. Lyding, IEEE Trans. Elec. Dev. **46**, 1812 (1999).

[10] I. C. Kizilyalli, G. C. Abeln, Z. Chen, J. Lee, G. Weber, B. Kotzias, S. Chetlur, J. W. Lyding, and K. Hess, IEEE Elect. Dev. Lett. **19**, 444 (1999).

[11] P. H. Lippel, R. J. Wilson, M. D. Miller, Ch. Wöll, and S. Chiang, Phys. Rev. Lett. **62**, 171 (1989).

[12] M. Kanai, T. Kawai, K. Motai, X. D. Wang, T. Hashizume, and T. Sakura, Surf. Sci. **329**, L619 (1995).

[13] J. K. Gimzewski, C. Joachim, R. R. Schlitter, V. Langlais, H. Tang, and I. Johannsen, Science **281**, 531 (1998).

[14] X. Yao, T. G. Ruskell, R. K. Workman, D. Sarid, and D. Chen, Surf. Sci. **367**, L85 (1996).

[15] J. Pollman, P. Krüger, and A. Mazur, J. Vac. Sci. Technol. **B5**, 945 (1987).

[16] M. C. Hersam, G. C. Abeln, and J. W. Lyding, Microelectron. Engng **47**, 235 (1999).

The influence of surfaces and interfaces on coherent phonons in semiconductors

C.-K. Sun, J. C. Liang

Department of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan, Republic of China

S. P. DenBaars

Department of Electrical and Computer Engineering, UCSB, Santa Barbara, CA 93106, U.S.A.

D. S. Kim, Y. D. Cho

Department of Physics, Seoul National University, Seoul 151-742, South Korea

G. D. Sanders, J. Simmons, C. J. Stanton

Department of Physics, University of Florida, Gainesville, FL 32611, U.S.A.

Experiments have shown that ultrafast optical excitation of semiconductors can produce oscillating changes in the optical properties of the material. The frequency of the oscillations in transmission or reflection usually matches one of the phonon modes, typically the $q = 0$ optical mode. These oscillations are known as *coherent phonons*. We discuss the role of surfaces and interfaces on the coherent phonon signal. We show that: (1) the coherent phonon signal can be used as a probe of the surface depletion field and (2) multiple interfaces as in a superlattice, can drastically alter the coherent phonon spectrum: screening of the modes in the superlattices is reduced and acoustic modes can now be excited.

© 2000 Academic Press

Key words: coherent-phonons, superlattices, ultrafast excitations.

1. Introduction

Femtosecond lasers have proven to be powerful tools for studying the dynamical behavior of photoexcited electrons in semiconductors. Recent experiments have shown that in addition, ultrafast excitation can produce oscillations in the optical transmission or reflection with a frequency matching one of the phonon modes, usually the $q = 0$ optical mode. These oscillations are known as *coherent phonons* [1–4].

In this paper, we discuss coherent phonons and the effect of surfaces and interfaces on the spectrum of oscillations.

The microscopic theory for non-polar materials is based on a two-band (electron and hole) model with the Hamiltonian given by [5, 6]

$$\hat{H}_{el} = \sum_{\mathbf{k},\alpha} \epsilon_{\alpha\mathbf{k}} c_{\alpha\mathbf{k}}^{\dagger} c_{\alpha\mathbf{k}} + \sum_{\mathbf{q}} \hbar_{\mathbf{q}} \omega_{\mathbf{q}} b_{\mathbf{q}}^{\dagger} b_{\mathbf{q}} + \sum_{\mathbf{k},\mathbf{q},\alpha} M_{\mathbf{k}\mathbf{q}} \left(b_{\mathbf{q}} + b_{-\mathbf{q}}^{\dagger} \right) c_{\alpha\mathbf{k}}^{\dagger} c_{\alpha\mathbf{k}+\mathbf{q}}. \tag{1}$$

The first term in eqn (1) represents the Bloch electrons/holes in the noninteracting crystal. Here c, $c^\dagger$ are the electron second quantization operators in $\mathbf{k}$-space, $\epsilon_{\alpha\mathbf{k}}$ is the energy dispersion in the conduction and valence bands ($\alpha = \{c, v\}$ indices the conduction or valence band). The second term represents the optical phonons. Here $\omega_\mathbf{q}$, $b^\dagger_\mathbf{q}$, $b_\mathbf{q}$ are the phonon dispersion and the phonon creation and annihilation operators. The electrons interact with the optical phonons through the third term. $M_{\mathbf{kq}}$ is the electron–phonon matrix element which typically might describe deformational coupling.

The *coherent amplitude* of the **q**th phonon mode is given by:

$$D_\mathbf{q} \equiv \langle b_\mathbf{q} \rangle + \langle b^\dagger_{-\mathbf{q}} \rangle \equiv B_\mathbf{q} + B^*_{-\mathbf{q}} \qquad (2)$$

and is proportional to Fourier components of the displacement.

One can obtain the equation of motion for the *coherent amplitude* by commuting eqn (2) with the Hamiltonian eqn (1):

$$\frac{\partial^2}{\partial t^2} D_\mathbf{q} + \omega_\mathbf{q}^2 D_\mathbf{q} = -2\omega_\mathbf{q} \sum_{\alpha,\mathbf{k}} M^\alpha_{\mathbf{kq}} n^\alpha_{\mathbf{k},\mathbf{k}+\mathbf{q}}. \qquad (3)$$

This is the equation of a forced oscillator. The right-hand side of eqn (3) depends on the Fourier transform of the electron density matrix and acts as a driving term. When electrons and holes are rapidly created during the laser pulse, this sets off the coherent oscillations. Note that since the laser light is macroscopic, usually the phonon mode that is coherently driven is the $\mathbf{q} \approx 0$ optical mode.

2. Polar materials

In polar materials such as GaAs, the coherent phonons are quite pronounced. The mechanism for generation is easily understood. This is illustrated in Fig. 1. A depletion field either from surface states or an externally applied field exists near the surface, as seen in Fig. 1A. This causes the polar lattice to distort in response to the field as shown in Fig. 1C. When electrons and holes are photoexcited by the ultrafast laser pulse, they screen out the depletion field and cause the lattice to relax back to its equilibrium position, Fig. 1B. This triggers oscillations not only in the lattice, but also the electron gas. Details of the interplay between the two are discussed in the paper of Kuznetsov [7, 8].

Two important effects emerge. (i) The amplitude of the coherent phonons depends on the strength of the surface depletion field. (ii) At high carrier excitation density, the photoexcited electrons screen the lattice mode and the frequency shifts from that of the LO mode to that of the TO mode.

3. Superlattices

In heterostructures and superlattices, the nature of the coherent phonons can change owing to the multiple interfaces [9]. One major effect that can occur is that the nature of the screening of the lattice modes by the photoexcited carriers can change dramatically. This results from the fact that in a bulk system, the plasmon is optical-like having a finite energy for $\mathbf{q} = 0$, while in a two-dimensional (2D) system, the plasmon energy goes to 0 as the wavevector goes to 0. The plasmon in a 2D system is unable to screen the LO mode independent of the excitation density. This is illustrated in Fig. 2. In Fig. 2A we show the coherent phonon oscillations in the time domain for bulk GaAs and a GaAs/Al$_{0.36}$Ga$_{0.64}$As MQW. The beating occurs between the LO and TO modes, originating from the fact that the density profile of the photoexcited carriers is not uniform. In Fig. 2B the Fourier transform in the frequency domain is plotted. One can see that at high carrier densities, the photoexcited electrons can not screen the LO mode as they can in bulk.

Further studies show that the screening of the LO mode in the superlattice depends in detail on the width and height of the barriers. A systematic study has been performed by Yee *et al.* [10]. These results show that
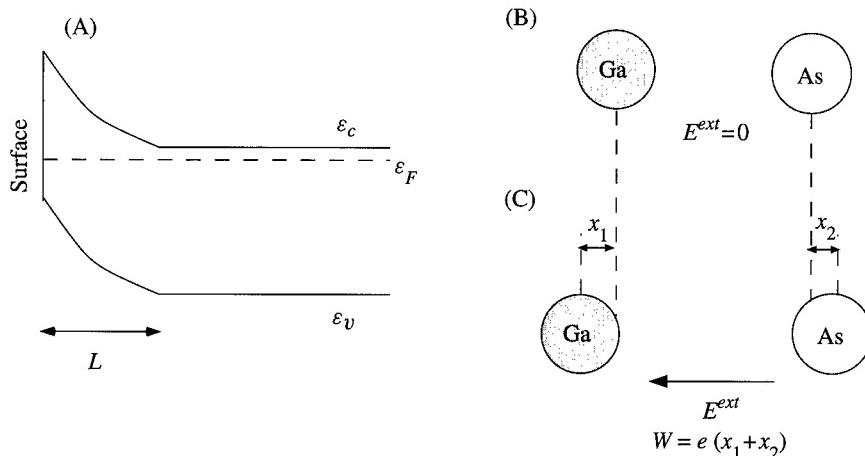
Fig. 1. The energy band diagram of a surface depletion region in GaAs is shown in A. In the depletion region, there are few carriers in a region of length L and large built in fields, E^{ext}. In the large electric field, the Ga and As atoms will move C, with respect to their positions with no electric field B, and create a lattice polarization W.
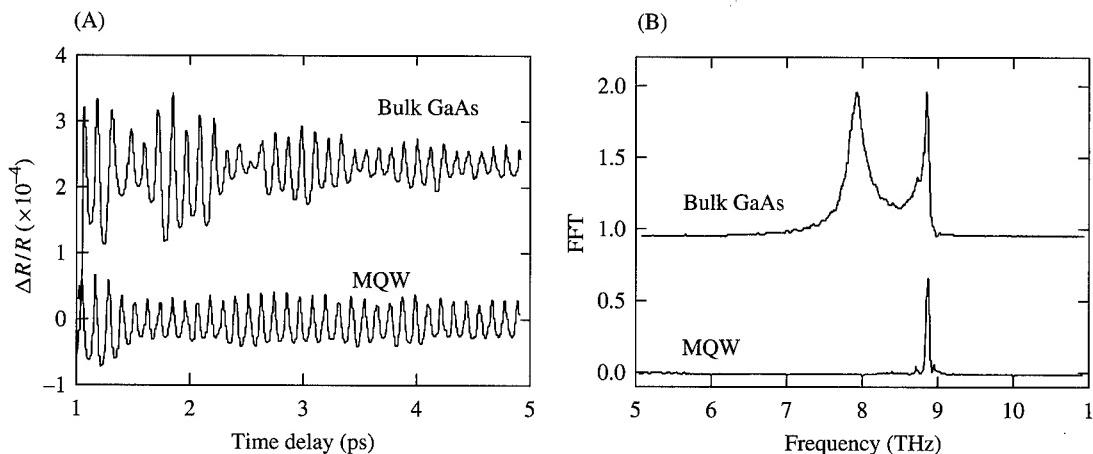
Fig. 2. A, Coherent optical phonon oscillations in the time domain for bulk GaAs and GaAs/Al$_{0.36}$Ga$_{0.64}$As MQW with $L_w = 15$ nm and $L_b = 5$ nm. B, The Fourier transform in the frequency domain. At high densities, the photoexcited electrons can not screen the LO mode in the MQW since the plasma dispersion varies from the bulk dispersion.

the density dependence of the frequency of oscillation can provide valuable information on the nature and quality of the interfaces in a superlattice.

More interesting behavior occurs in strained layer superlattices such as the InGaN/GaN systems. Here, owing to the wurtzite structure, strain in the superlattice leads to large piezoelectric fields. The large fields lead to strong oscillations, as shown in Fig. 3. These oscillations are so large that they can be seen in transmission and are much stronger than those seen in Fig. 2 in reflection.

Even more interesting is the fact that the oscillations are much slower than in bulk GaAs as can be seen by comparing Figs 2 and 3. In fact, the oscillations are now at the *acoustic* frequency. The mechanism for the generation of the acoustic phonons can be seen from eqn (3). Since absorption in the superlattice occurs only within the wells, the laser pulse can produce an electron density that has a Fourier component determined by
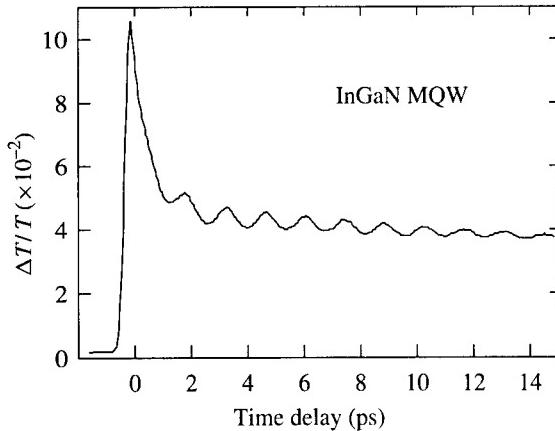
Fig. 3. Transient transmission changes versus probe delay for a 50/43 Å InGaN/GaN MQW and a laser wavelength of 365 nm. Huge coherent *acoustic* phonon oscillations are now observed.

the superlattice period. This has been investigated in more detail by Sun *et al.* [11]. This result shows that the superlattice can significantly change the coherent phonon spectrum and suggests that band-engineering heterostuctures can lead to tailored coherent oscillations.

4. Conclusions

We have shown that surfaces and interfaces can alter the properties of the coherent phonons created under ultrafast laser excitation. First, the strength of the oscillations is sensitive to the magnitude of the surface depletion field. Second, interfaces can drastically modify the spectrum: screening of the modes in superlattices for high-density excitation is reduced and acoustic modes can now be excited. This suggests that coherent phonons might be used as sensitive probes of surfaces and interfaces.

References

[1] G. C. Cho, W. Kutt, and H. Kurz, Phys. Rev. Lett. **65**, 764 (1990).
[2] W. Kutt, G. C. Cho, T. Pfeifer, and H. Kurz, Semicond. Sci. Technol. **B77**, 7 (1992).
[3] W. Albrecht, Th. Kruse, and H. Kurz, Phys. Rev. Lett. **69**, 1451 (1992).
[4] W. A. Kutt, W. Albrecht, and H. Kurz, IEEE J. Quantum Electronics **QE-28**, 2434 (1992).
[5] A. V. Kuznetsov and C. J. Stanton, Phys. Rev. Lett. **73**, 3243 (1994).
[6] R. Merlin, Solid State Commun. **102**, 207 (1997).
[7] A. V. Kuznetsov and C. J. Stanton, Phys. Rev. **B52**, 7555 (1995).
[8] A. V. Kuznetsov and C. J. Stanton, SPIE **3277**, 28 (1998).
[9] T. Dekorsey, A. M. T. Kim, G. C. Cho, H. Kurz, A. V. Kuznetsov, and A. Forster, Phys. Rev. **B53**, 1531 (1996).
[10] K. J. Yee, D. S. Kim, T. Dekorsey, and Y. S. Lim, Phys. Rev. **B60**, R8513 (1999).
[11] C.-K. Sun, J.-C. Liang, C. J. Stanton, A. Abare, L. Coldren, and S. DenBaars, Appl. Phys. Lett. **1249**, 75 (1999).

Silicon-based nanoelectronics and nanoelectromechanics

A. Tilke, A. Erbe, L. Pescini, H. Krömmer, R. H. Blick, H. Lorenz, J. P. Kotthaus[†]

Center for NanoScience and Sektion Physik, Ludwig-Maximilians-Universität, Geschwister-Scholl-Platz 1, 80539 Munich, Germany

We demonstrate Coulomb blockade oscillations in different single-electron devices in Silicon-On-Insulator (SOI) films up to temperatures of 300 K. The layer sequence in SOI allows the underetching of these devices in order to realize suspended, highly doped silicon nanostructures. Similar suspended silicon beams are fabricated to form novel nanomechanical resonators that can be excited at radio frequencies up to about 300 MHz. Controlling the vibration frequency by a side-gate voltage, these resonators allow charge detection with a sensitivity of 0.1 $e/\sqrt{Hz}$, comparable to that of cryogenic single-electron devices.

© 2000 Academic Press

Key words: single-electron devices, nanomechanics.

1. Introduction

High-resolution low-energy electron-beam lithography using the negative electron resist calixarene gives us the definition of device dimensions down to 10 nm for the fabrication of nanostructures in thin Silicon on Insulator (SOI) films [1]. We realize both highly doped nanowires [2] to study Coulomb blockade (CB) in the classical metallic regime as well as quantum dots defined in inversion layers of metal-oxide field effect transistors (MOSFETs) [3] to observe non-periodic CB oscillations strongly affected by spatial quantization. Underetching of these nanostructures in the SOI-system allows to realize suspended single electron transistors (SuSETs) [4] and, by metallization of these suspended beams, also a novel kind of nanomechanical resonators [5].

2. Quasi-metallic Coulomb blockade oscillations

We combine the well-established SOI-technology with the fabrication of quasi-metallic narrow Si-wires by an extremely high doping of our SOI-films [2]. In many cases these act as metallic single-electron transistors (SETs) and offer many similar charging states and hence a broad range of operating points. Our samples have a nominal doping level of 10^{21} cm^{-3} realized by ion-implantation of arsenic. This leads to a mean distance of only 1 nm between the dopants. Figure 1 depicts the conductance of a 50 nm wide nanowire at a temperature of 4.2 K as a function of the applied topgate-voltage. CB oscillations with a clear periodicity are observed, indicating the metallic nature of the SET. From the slope of the CB diamond we determine the

---

[†]E-mail: jorg.kotthaus@physik.uni-muenchen.de
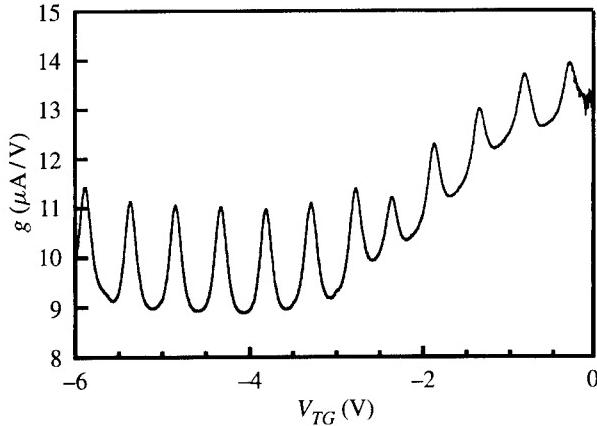
Fig. 1. Coulomb blockade oscillations in a quasi-metallic nanowire showing almost perfect periodicity.

charging energy in the dot presented to be $E_C = 17.5$ meV. Presumably, in this quasi-metallic nanostructure, the pattern-dependent two-dimensional oxidation of narrow wires in combination with edge roughness and the pile-up effect of As-dopants during dry oxidation [6] lead to the observed SET-behavior.

3. Coulomb blockade in inversion-channel quantum dots

In contrast to the use of highly doped SOI-films, embedding lithographically defined SET-structures into the inversion channel of SOI-field effect transistors allows us to observe CB oscillations up to very high temperatures [2, 7]. Since the tunneling barriers are defined by geometry, the controllability of these devices is greatly enhanced in comparison with random effects found in highly doped nanowires. We prepared geometrically defined quantum dot structures embedded in SOI-MOSFETs with lateral dimensions below 20 nm [3]. In Fig. 2A the CB oscillations in a dot with an estimated lithographical diameter of 15 nm are presented. A strong influence of spatial quantization inside the dot leads to a deviation of the classical periodicity of the CB oscillations. From the slope of the CB diamond and the mean spacing between two adjacent conductance oscillations one can deduce the charging energy to $E_C = 56$ meV. In Fig. 2B the source–drain I–V characteristics are shown for various temperatures. Figure 2B also shows the I–V characteristics as well as the source–drain conductance at 300 K. Clear CB operation at room temperature is reflected in the minimum of the conductance g around zero bias.

4. Suspended single-electron structures

Underetching the SiO_2-spacer of the lithographically defined SOI-nanostructures leads to a suspension of the single-electron devices and therefore to a thermal decoupling from the silicon substrate. The buried oxide is removed locally in buffered hydrofluoric acid and the sample subsequently rinsed in water and isopropanol [4]. Due to random dopant fluctuations, single-electron effects become visible in these wires at low temperatures.

5. Nanoresonators

Nonlinear micromechanical resonators have been shown to be very efficient tools for charge detection [8]. Scaling down these resonators to dimensions in the 100 nm range leads to even higher accuracies [5]. Cor-
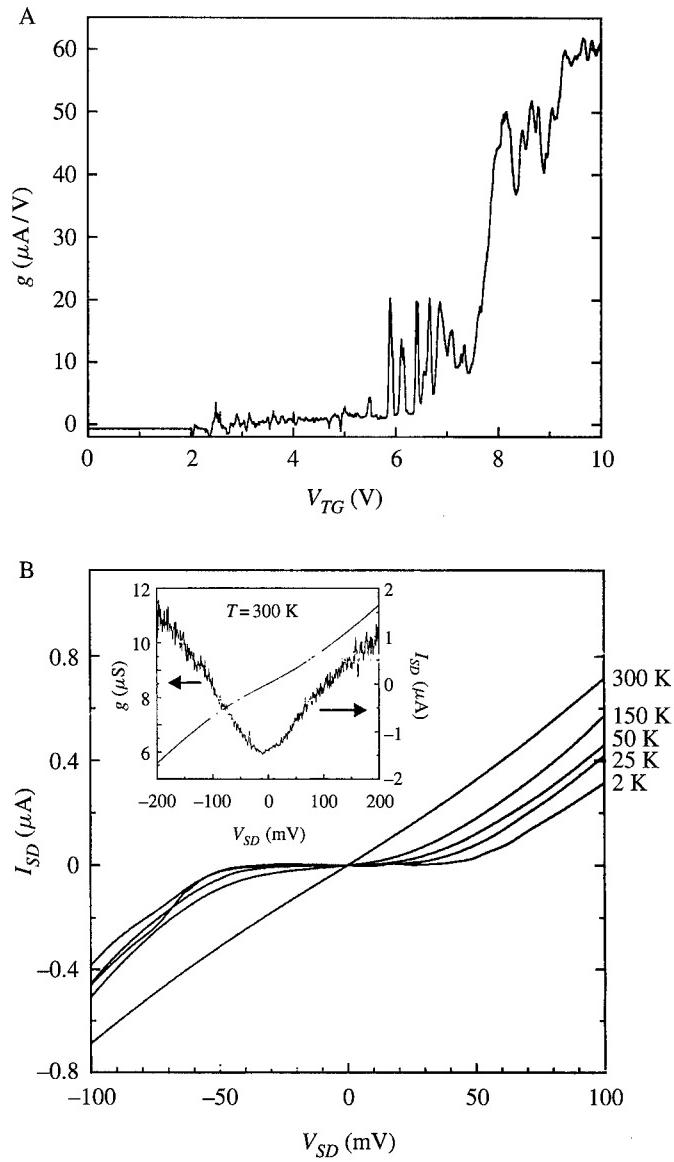
Fig. 2. A, Dependence of the conductance g of an inversion quantum dot on gate voltage V_{TG}. Coulomb blockade is observed near threshold voltage. B, Temperature dependence of the $I_{SD}-V_{SD}$ trace of the dot at a top gate voltage of $V_{TG} = 4.5$ V. At 300 K clear single-electron effects remain visible.

respondingly, the frequencies are increased to the rf-regime, enhancing the speed of charge detection significantly. Our resonators are machined out of SOI material similar to the doped suspended structures. A SEM-micrograph of a final structure is shown in the inset of Fig. 3. This resonator has a length of almost 3 μm, width and thickness being 200 nm. This particular resonator is a straight beam of silicon covered with a thin layer of gold (50 nm). The two gates on the left and right sides of the beam can be biased capacitively to detune the resonators eigenfrequency. They can also be used to drive the resonator by applying an rf-voltage. The beam is cooled down to 4.2 K and placed in a magnetic field perpendicular to the beam's
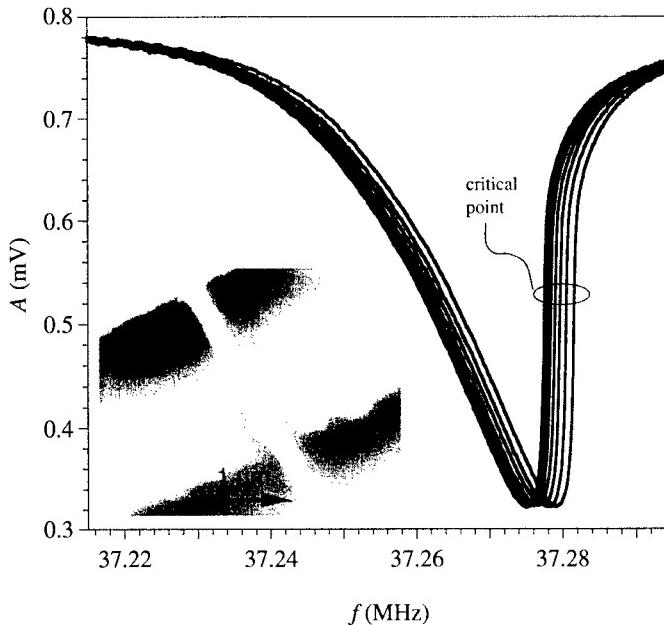
Fig. 3. Resonance curves of the reflected rf power of the suspended wire shown in the inset. The resonator is driven into the nonlinear regime by the Lorentz force in a B-field of 12 T applied perpendicularly to the rf current. Different curves are measured at different values of the voltage on one of the sidegates. Inset: SEM micrograph of the structure.

direction. If a current is driven through the beam, the magnetic field induces a Lorentz force, which in turn sets the resonator into motion.

If the resonator is driven by a small ac-current, the resulting resonance curve is fully symmetrical. When increasing the ac current, the resonator is driven into nonlinear response. The motion of the center of mass is then properly described by the well-known Duffing equation, i.e. the eigenfrequency shifts and the resonance curve becomes asymmetrical [9]. If one of the gates is charged by applying a voltage V with respect to the wire, the eigenfrequency changes quadratically with V and hence an additional term proportional to V^2 appears in the Duffing equation. This change can best be seen, if the resonator is driven in the transition regime between linear and nonlinear behaviour. Then the resonance curve exhibits a region of infinite derivative (critical point). The variation of the position of this region can be monitored in dependence of the gate's potential.

This behavior is demonstrated in Fig. 3. The shift in the resonance frequency can be best resolved at the point of large derivative, which is indicated as critical point. Due to the quadratic dependence of the resonance frequency on the applied voltage the charge resolution is increased when increasing the gate voltage. At a gate voltage of 4 V, we find a resolution of 0.1 $e/\sqrt{Hz}$. The resolution is thus comparable to common SET charge detectors working at much lower temperatures.

References

[1] A. Tilke, M. Vogel, F. Simmel, A. Kriele, R. H. Blick, H. Lorenz, D. A. Wharam, and J. P. Kotthaus, J. Vac. Sci. Technol. B **17**, 1594 (1999).
[2] A. Tilke, R. H. Blick, H. Lorenz, J. P. Kotthaus, and D. A. Wharam, Appl. Phys. Lett. **75**, 3704 (1999).
[3] A. Tilke, L. Pescini, R. H. Blick, H. Lorenz, and J. P. Kotthaus, submitted to Appl. Phys. A, (1999).
[4] L. Pescini, A. Tilke, R. H. Blick, H. Lorenz, J. P. Kotthaus, W. Eberhardt, and D. Kern, Nanotechnology **10**, 418 (1999).

[5] H. Krömmer, A. Erbe, A. Tilke, S. Manus, and R. H. Blick, Europhys. Lett. **50**, 10 (1999).

[6] A. S. Grove, O. Leistiko Jr., and C. T. Sah, J. Appl. Phys. **35**, 2695 (1964).

[7] E. Leobandung, L. Guo, Y. Wang, and S. Y. Chou, Appl. Phys. Lett. **67**, 938 (1995).

[8] A. N. Cleland and M. L. Roukes, Nature **392**, 160 (1998).

[9] U. Parlitz and W. Lauterborn, Phys. Lett. **107A**, 351 (1985).

Resistively coupled single-electron transistor using tunnel gate resistor

FUJIO WAKAYA[†], SHINYA NAKAMICHI, KENJI GAMO[†]

Department of Physical Science, Graduate School of Engineering Science, Osaka University,
1–3 Machikaneyama-cho, Toyonaka, Osaka 560-8531, Japan

SHUNJI MANDAI

Department of Electrical Engineering, Faculty of Engineering Science, Osaka University,
1–3 Machikaneyama-cho, Toyonaka, Osaka 560-8531, Japan

SHUICHI IWABUCHI

Department of Physics, Faculty of Science, Nara Women's University, Kitauoya-Nishimachi, Nara
630-8506, Japan

A resistively coupled single-electron transistor (R-SET) was fabricated using a modulation-doped heterostructure and metal Schottky gates, and measured at low temperature. Currents of R-SET with tunnel gate resistor were calculated using the orthodox theory. It is shown that the R-SET with tunnel gate resistor has quite similar properties to the originally proposed R-SET.

© 2000 Academic Press

Key words: R-SET, C-SET, Coulomb blockade, Coulomb diamond.

1. Introduction

Since single-electron transistors (SETs) were proposed by Likharev [1, 2], the field of Coulomb blockade and single-electron tunneling has been attracting researchers of both fundamental and applied science [3]. Likharev proposed two types of SET in his pioneering work [1, 2]. One is capacitively coupled SET (C-SET) and the other is resistively coupled SET (R-SET), both of which are depicted in Fig. 1 with the Coulomb-blockade regions. Although R-SETs have advantages of high-voltage gain and background-charge-free operation, C-SETs and C-SET-based devices have been intensively studied so far, while a few works [4–6] on R-SET can be found in the literature. The purpose of the present work is fabrication and characterization of the R-SET.

2. Experiments

We fabricated an SET in two-dimensional electron gas (2DEG) in GaAs/AlGaAs heterostructure using metal Schottky gates on the surface (Fig. 2). A modulation-doped heterostructure was chemically etched

---

[†]Also at: Research Center for Materials Science at Extreme Conditions, Osaka University, 1–3 Machikaneyama-cho, Toyonaka, Osaka 560-8531, Japan
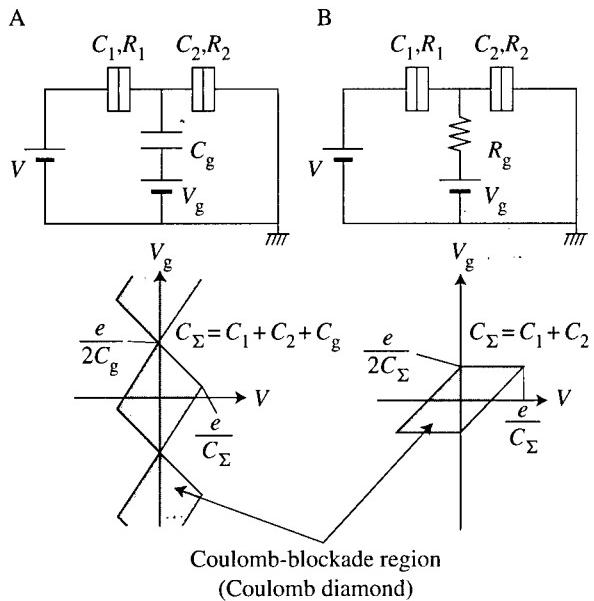
Fig. 1. Schematic diagrams of A, C-SET and B, R-SET with their Coulomb diamonds.
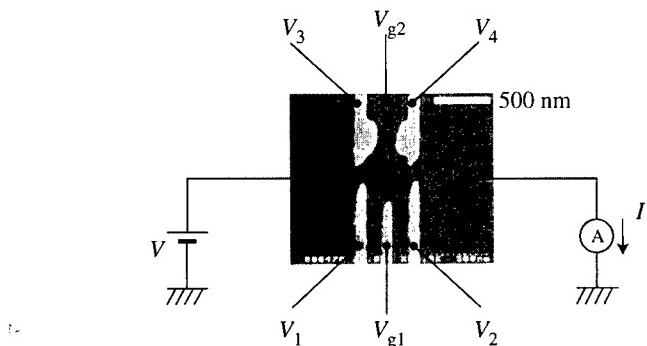
Fig. 2. SEM micrograph of the fabricated device. The measurement set-up is also shown.

to define a Hall-bar structure. Au/Ni/AuGe was deposited and annealed to make ohmic contacts. Electron-beam lithography and a lift-off technique of a deposited AuPd thin film were used to define the fine Schottky gates. The Coulomb island and the tunnel junctions were formed by depleting the 2DEG using negative gate voltages on the Schottky gates. The gate resistor, which is a key component of the R-SET, was also realized by depleting the 2DEG between the gates whose voltages are V_3 and V_4 which are defined in Fig. 2. The current–voltage (I–V) measurements were performed at $\simeq 30$ mK using a dilution refrigerator. The measurement set-up is also shown in Fig. 2.
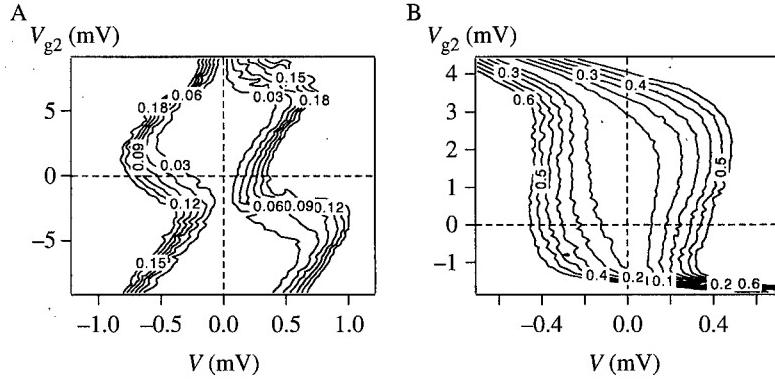
Fig. 3. Observed Coulomb diamonds. A, $|I| = 0.03, 0.06, \ldots, 0.18$ nA. $V_1 = -1.194$ V, $V_2 = -1.324$ V, $V_3 = -0.444$ V, $V_4 = -0.444$ V, $V_{g1} = -0.52$ V. B, $|I| = 0.1, 0.2, \ldots, 0.6$ nA. $V_1 = -1.228$ V, $V_2 = -1.333$ V, $V_3 = -0.415$ V, $V_4 = -0.415$ V, $V_{g1} = -0.52$ V.

3. Results and discussion

When the voltages V_3 and V_4, which are defined in Fig. 2, are sufficiently negative, the gate resistance R_g should actually be infinity. In such a case, the device should show C-SET characteristics. Figure 3A shows the observed Coulomb diamonds with $V_1 = -1.194$ V, $V_2 = -1.324$ V, $V_3 = -0.444$ V, $V_4 = -0.444$ V and $V_{g1} = -0.52$ V. The observed C-SET-type diamond suggests that R_g can be recognized to be ∞. Figure 3B shows the observed Coulomb diamond with V_3 and V_4 more positive. V_1 and V_2 were also changed to be -1.228 V and 1.333 V, respectively, in order to keep the source and drain point contacts closed. Using these parameters, only one Coulomb diamond was observed as shown in Fig. 3B. In this case, the finite current was observed at $V = 0$ when $V_{g2} \gtrsim 3$ mV or $V_{g2} \lesssim -1$ mV. This current flew between the gate and source electrodes, i.e. R_g is not ∞ in this case. However, the center of the observed diamond was not located at the origin of the V–V_{g2} plane, while it should be at the origin according to the original theory. Moreover, the shape of the observed Coulomb diamond was not exactly the same as the theoretically predicted one.

The shift of diamond is usually observed at a C-SET with offset charge at the island, which is connected to the surrounding electrodes through the tunnel junctions. The observed shift of the diamond (Fig. 3B), therefore, strongly suggests that the fabricated device was affected by the offset charge at the island. This means that the island is surrounded by only tunnel junctions, i.e. the fabricated device has the tunnel gate resistor.

In order to examine what happens in the R-SET with a tunnel gate resistor whose tunnel capacitance and resistance are C_g and R_g, respectively, we calculated I–V characteristics using so-called orthodox theory [3]. Figure 4 shows the typical calculated Coulomb diamonds. The parameters used are shown in the caption, where $U = e^2/(2C_\Sigma)$ is the charging energy and Q_0 is the offset charge. As shown in Fig. 4A the device with tunnel gate resistor shows C-SET-type Coulomb diamonds when R_g is sufficiently larger than R_1 and R_2. This should be the case of Fig. 3A in the experiment. Figure 4B suggests that the R-SET with tunnel gate resistor can show a quite similar Coulomb diamond to the originally proposed R-SET. One of the differences is that the former is affected by the offset charge as shown in the figure. Therefore, the shift of diamond along the V_g-axis which is seen in Fig. 3B can be explained if the fabricated device has the tunnel gate resistor.

As shown in the SEM photograph (Fig. 2), the shape of the electrodes whose voltage were V_3 and V_4 is not very straight, somewhat round. The depletion regions of the two gates (V_3 and V_4), therefore, should touch each other first at the narrowest part, where the tunnel junction formed. This should be the reason why the fabricated device has the tunnel gate resistor.
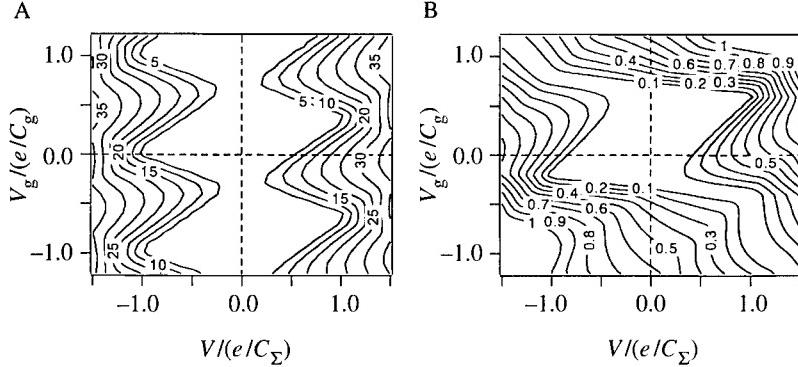
Fig. 4. Calculated Coulomb diamonds. A, $|I/(e/R_\Sigma C_\Sigma)| = 5, 10, \ldots, 35.$ $R_g/R_1 = R_g/R_2 = 100, C_g/C_1 - C_g/C_2 = 1, U/(kT) = 100, Q_0/e = -0.2.$ B, $|I/(e/R_\Sigma C_\Sigma)| = 0.1, 0.2, \ldots, 1.0.$ $R_g/R_1 = R_g/R_2 = 1, C_g/C_1 = C_g/C_2 = 10, U/(kT) = 100, Q_0/e = -0.2.$

4. Summary

We fabricated an R-SET using a modulation-doped heterostructure and metal Schottky gates and measured I–V characteristics. The observed Coulomb diamond did not fully satisfy the originally predicted features. In order to discuss the origin of the deviation, the numerical calculations were performed. Comparing the experimental and numerical results, the fabricated device should have a tunnel gate resistor. The numerical results show that the R-SET with tunnel gate resistor has quite similar properties to the originally proposed R-SET.

References

[1] K. K. Likharev, IEEE Trans. Magn. **23**, 1142 (1987).

[2] K. K. Likharev, IBM J. Res. Dev. **32**, 144 (1988).

[3] See for example, *Single Charge Tunneling*, edited by H. Grabert and M. H. Devoret (Plenum Press, New York, 1992).

[4] P. Delsing, T. Claeson, G. S. Kazacha, L. S. Kuzmin, and K. K. Likharev, IEEE Trans. Magn. **27**, 2581 (1991).

[5] N. Yoshikawa, Y. Jinguu, H. Ishibashi, and M. Sugahara, Jpn. J. Appl. Phys. **35**, 1140 (1996).

[6] F. Wakaya, K. Kitamura, S. Iwabuchi, and K. Gamo, Jpn. J. Appl. Phys. **38**, 2470 (1999).

Multiple-valued logic devices using single-electron circuits

TAKASHI YAMADA, YOSHIHITO AMEMIYA

The Department of Electrical Engineering, Hokkaido University, Kita 13, Nishi 8, 060-8628, Sapporo, Japan

Multiple-valued logic devices can be constructed compactly by utilizing quantized behavior of single-electron circuits. As an example, a single-electron multiple-valued Hopfield network solving optimization problems is designed. Computer simulation shows that the network can successfully converge to its optimal state that represents the solution to the problem.

© 2000 Academic Press

Key words: single electron, multiple valued, Hopfield network.

1. Introduction

One of the challenges in microelectronics is the development of novel electronic devices that can perform functional computing by utilizing inherent properties of quantum phenomena. We here propose one such computation device, namely a *multiple-valued logic device using single-electron circuits.*

The multiple-valued logic is a way of implementing digital operations by using a multiple set of logical values $\{0, 1, 2, 3, \ldots\}$ instead of a binary set $\{0, 1\}$. The logic process of the multiple-valued logic is much more sophisticated than that of binary logics, so the multiple-valued logic is expected to be more powerful for implementing digital functions with a smaller number of devices. But in practice, it has been unsuccessful to construct multiple-valued logic LSIs because most of multiple-valued logic functions are hard to implement using CMOS circuits.

To overcome this problem, we here propose an idea that multiple-valued logic systems can be constructed into a compact circuit by using the single-electron circuit technology. The single-electron circuit is a quantum electronic circuit based on the Coulomb blockade effect in electron tunneling. A conspicuous property of the single-electron circuit is that the circuit shows 'quantized behavior' in its operation; i.e. the variation of the charge on each node of a single-electron circuit is quantized in units of the elementary charge because the node charge is changed only through electron tunnelings. By utilizing this, we can construct various circuits for multiple-valued logic operations. As an example, implementation of the multiple-valued Hopfield network is discussed in the following.

2. The multiple-valued Hopfield network

The multiple-valued Hopfield network is a computation device for solving combinatorial optimization problems with multiple-valued variables. The concept of the network is illustrated in Fig. 1. The network consists of many neurons and connections. The output of each neuron feeds back into inputs of other neurons.
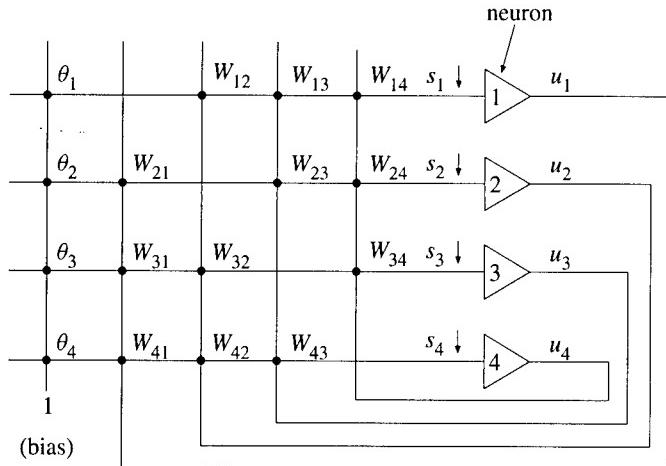
Fig. 1. Hopfield network.
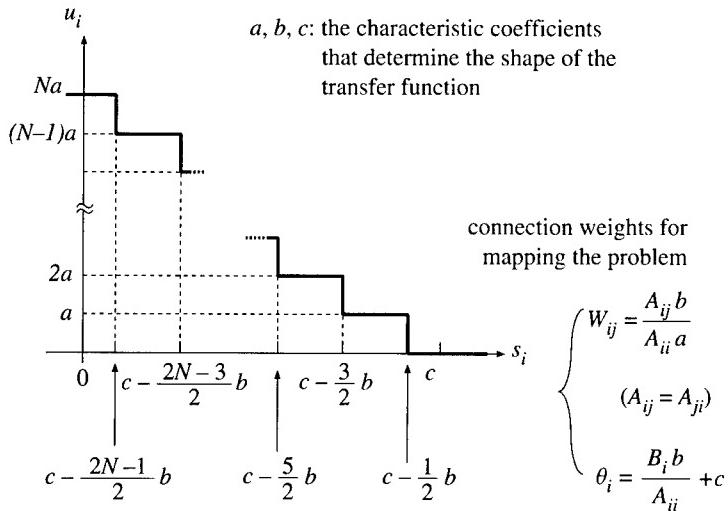
Fig. 2. Staircase transfer function.

Denoted by W_{ij} is the connection weight to neuron i from neuron j, and θ_i is the bias-connection weight to neuron i from a bias that fixed at the value of 1, and u_i is the output of neuron i. A set of neuron outputs $(u_1, u_2, \ldots)$ is called the state of the network. In this network, each neuron i takes a weighted sum s_i of its inputs $(s_i = \sum_j W_{ij} u_j + \theta_i)$ and generates the corresponding output u_i according to a given staircase transfer function. All neurons operate in parallel to update their outputs continuously, and the network converges to an optimal state through the updating process.

The structure of combinatorial optimization problems can be mapped into the structure of the network by deciding the connection weights between neurons. As an example, we here take up the quadratic integer-programming problem: Given a set of coefficients A_{ij} and B_i $(A_{ij} = A_{ji}, A_{ii} \neq 0; i, j = 1, 2, 3, \ldots, n)$,
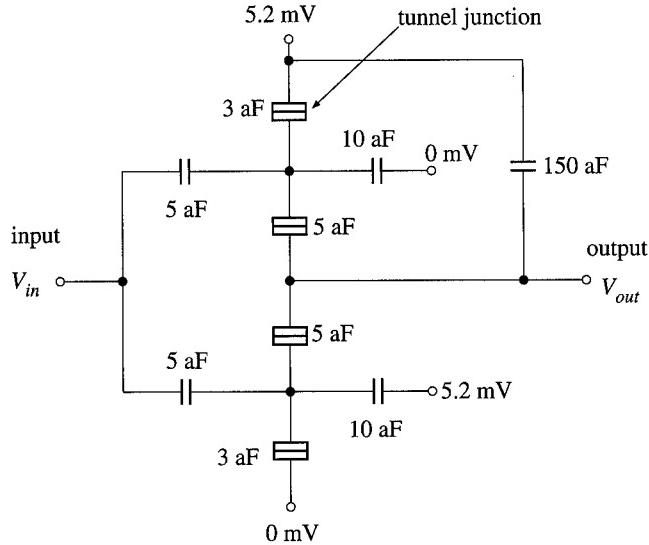
Fig. 3. Single-electron neuron circuit.

minimize objective function

$$\frac{1}{2}\sum_{i}^{n}\sum_{j}^{n} A_{ij}x_i x_j + \sum_{i}^{n} B_i x_i, \quad \text{where } x_i \in 0, 1, 2, \ldots, N \ (i = 1, 2, \ldots, n). \tag{1}$$

To solve this problem, we prepare a multiple-valued Hopfield network such that the output u_i of each neuron represents each problem variable x_i; we prepare neurons with the staircase transfer function illustrated in Fig. 2 and represent problem variables x_i by $x_i \equiv u_i/a$. The connection weights required for mapping the problem are given in Fig. 2. In problem solving, we set the network in an initial state (any state will do), then allow it to change its state without restraint. After some transition time the network converges to a final state. If convergence has been successful, the solution to the problem is obtained from the final state of the network (we will not discuss the local-minimum effect here).

3. Constructing a multiple-valued neuron using single-electron circuits

The single-electron circuit is an electronic circuit consisting of tunnel junctions and capacitors that is designed for manipulating electronic functions by controlling the transport of individual electrons. The internal state of the circuit is determined by the configuration of its electrons (i.e. the pattern in which the excess electrons are distributed among the nodes of the circuit). The circuit changes its electron configuration through electron tunnelings in response to the input and, thereby, changes its output voltage as a function of inputs.

Our purpose is to construct a single-electron neuron circuit that changes the charge on its output node (therefore its output voltage) according to a staircase function of the input voltage. For this purpose, we take *Tucker's single-electron inverter* and modify its circuit parameters to create a staircase transfer function (for the details of Tucker's original circuit, see [1]). The circuit we use for the neuron devices is illustrated in Fig. 3 together with a sample set of device parameters. The corresponding transfer characteristic are illustrated in Fig. 4. The number of steps in the transfer characteristic can be controlled by designing the circuit parameters.
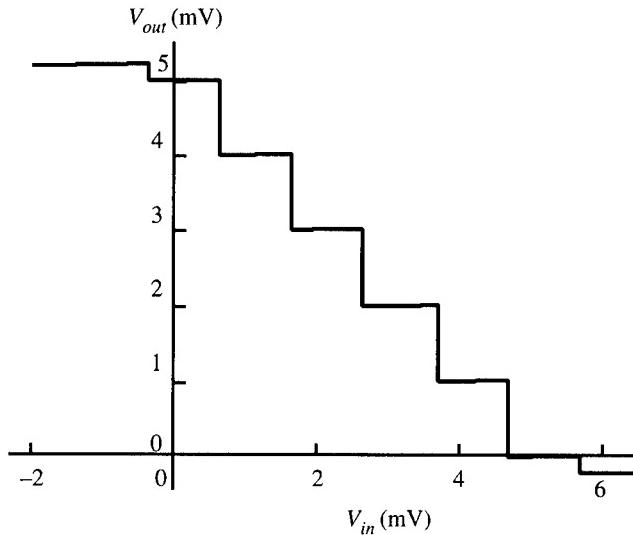
Fig. 4. Staircase transfer characteristic.
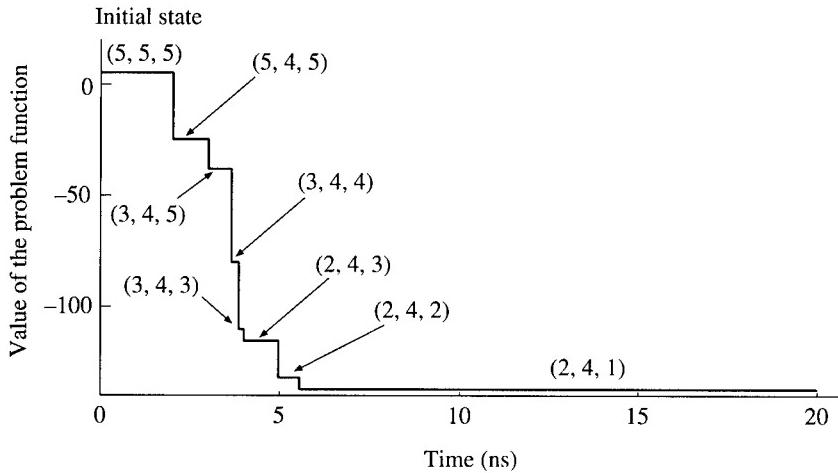
Fig. 5. State transition of the network (simulation).

4. Operation of a multiple-valued Hopfield network

The multiple-valued Hopfield network can be constructed by combining the developed neuron circuits into a network. We illustrate here a sample network that solves an instance of the quadratic integer-programming problems.

Consider the following problem:
Minimize

$$3x_1^2 + 6x_2^2 + 6x_3^2 + 4x_1x_2 - 2x_2x_3 + x_3x_1 - 29x_1 - 53x_2 - 7x_3, \qquad \text{where } x_i \in \{0, 1, 2, 3, 4, 5\}. \quad (2)$$

To solve this problem instance, we prepare a network with three neuron to represent problem variables x_i by output u_i of i-th neurons ($i = 1, 2, 3$). We used the neuron circuit with the transfer characteristic

illustrated in Fig. 4 and determined the value of connections for mapping the given problem. For problem solving, it is essential that, starting with a given initial state, the network circuit should converge to its global minimum energy states. To observe the behavior of the sample network, we simulated the state transition of the network.

The results of the simulation are illustrated in Fig. 5 in which the state of the sample network is expressed by a normalized set of three neuron outputs $(u_1/a, u_2/a, u_3/a)$, where u_i/a represent problem variables x_i. The circuit was initially set at state (5, 5, 5), then it was allowed to change its state without restraint. After some transition time, the circuit stabilized in the final state (2, 4, 1) that corresponds to the solution to the problem. We repeated the same trial many times and confirmed that every trial resulted in successful convergence to state (2, 4, 1).

5. Conclusion

We proposed that multiple-valued logic devices can be constructed into a compact circuit by using single-electron circuit technology. To present a practical form of our idea, we designed a single-electron multiple-valued Hopfield network and confirmed that the network can successfully converge to its optimal state representing the solution to the problem. Our results show that multiple-valued LSIs can be fabricated by using single-electron circuit technology.

References

[1] J. R. Tucker, J. Appl. Phys. **72**, 4399 (1992).

Quantum-dot structures measuring Hamming distance for associative memories

TAKASHI MORIE, TOMOHIRO MATSUURA, SATOSHI MIYATA, TOSHIO YAMANAKA, MAKOTO NAGATA, ATSUSHI IWATA

Faculty of Engineering, Hiroshima University, Higashi-Hiroshima, 739-8527 Japan

Two types of quantum-dot circuits measuring a Hamming distance using the Coulomb repulsion effect are proposed and analysed. They have structures where a quantum-dot array is arranged on a gate electrode of an ultrasmall MOSFET. The device parameters for successful operation are clarified from Monte Carlo simulation.

© 2000 Academic Press

Key words: quantum dots, single-electron devices, Coulomb repulsion, Hamming distance, associative memory.

1. Introduction

Single-electron devices have intrinsic difficulties from a viewpoint of system applications: slow operation speed and insufficient reliability [1] due to stochastic tunneling events and serious background charge sensitivity [2]. Thus, the conventional multi-stage digital circuit architecture is not suitable for single-electron devices.

In our strategy for constructing single-electron circuits, single-electron devices are used for repeatable circuits with a few logic-stages, and ultrasmall CMOS devices are used for multi-stage logic circuits. The above-mentioned difficulties are overcome by massively parallel operation and redundant architecture by virtue of very large packing density and very low-power dissipation of single-electron devices.

In this paper, as an example of such circuits, we propose two functional circuits using the Coulomb repulsion effect between quantum dots. These circuits measure the Hamming distance, the number of the unmatched bits between two digital data, which are a core circuit for associative memories. Although such circuits using single-electron transistors have already been proposed [3,4], the circuits using quantum dots have advantages of achieving static operation and robustness to fluctuation of device parameters such as the tunneling resistance.

2. Associative memory architecture using quantum dots and ultrasmall CMOS devices

The associative memory assumed here compares the input pattern with the stored patterns, and extracts the pattern most similar to the input. We assume that each pattern consists of N bit binary data, which is referred to as a *word*, and the stored data consist of M words. The similarity between digital data is usually measured by a Hamming distance.
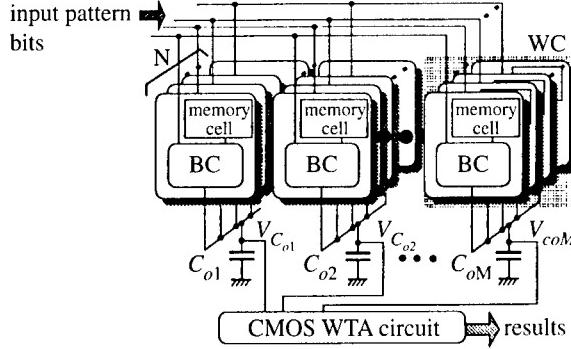
Fig. 1. Associative memory architecture.
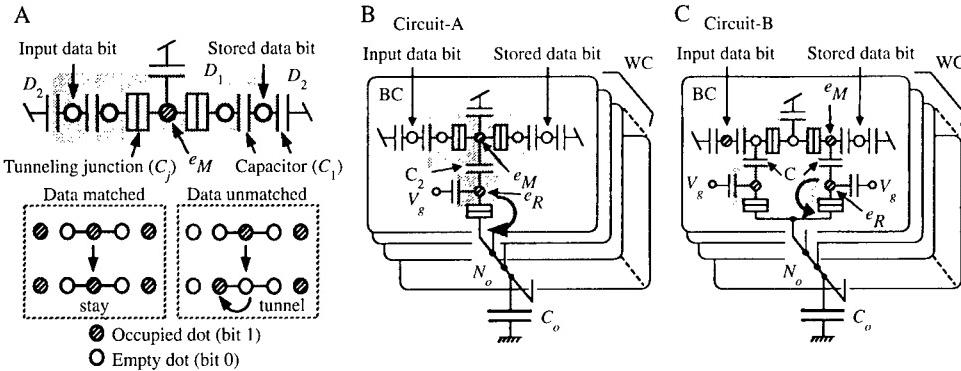
Fig. 2. Principle of bit-comparison using Coulomb repulsion effect (A), and two word-comparator circuits (B, C).

As shown in Fig. 1, bit-level comparison between the input data and each stored data is performed by each bit-comparator (BC) in a word-comparator (WC) in parallel. The WC calculates the Hamming distance between the input data and the stored data by collecting the results of the BCs; that is, the comparison result is expressed as the number of electrons released from the BCs, and the electrons are collected at the capacitor C_{oi}, $i = 1, \ldots, M$. All WC's results are fed into the winner-take-all (WTA) circuit, and it extracts the stored word data that has the shortest Hamming distance.

We assume that the WTA circuit is constructed using CMOS devices, and propose two quantum-dot circuits for the WC in the next section.

3. Quantum-dot circuits measuring Hamming distance

Let us assume a string of quantum dots as shown in Fig. 2A, put an electron, e_M, at one of the three dots D_1, and represent a bit (0 or 1) of the input and stored data by whether an electron is put at each end dot D_2 or not. When the corresponding bits of both data are matched, because Coulomb repulsion is symmetric, electron e_M is stabilized at the center; otherwise it is off-center. In order to detect the position of electron e_M, two detection circuits are proposed: circuit-A that detects the center position and circuit-B that detects the off-center position as shown in Fig. 2B and C. By the Coulomb repulsion effect, the bit matching result reflects whether electron e_R tunnels to node N_o. Consequently, electrons whose number is equal to that of
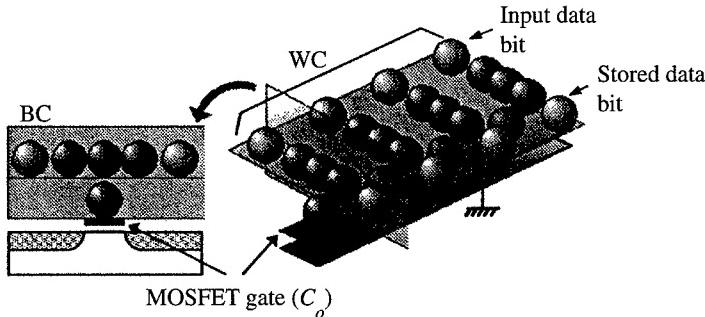
Fig. 3. Three-dimensional structure image of a word-comparator (circuit-A).

the matched bits (circuit-A) or that of the unmatched bits (circuit-B) are accumulated in C_o. To ensure the stabilizing processes, control voltage V_g are used.

A 3D arrangement of quantum dots realizing circuit-A is shown in Fig. 3, where common terminals to the ground and control voltages are omitted. The capacitance C_o corresponds to the gate capacitance of an ultrasmall CMOS transistor.

In this architecture, if it is difficult to represent one bit by one quantum dot, we can use a redundant architecture where plural BCs represent one data bit. Such an architecture based on a majority decision principle has the advantage of robustness against effects of background charge.

4. Circuit simulation results

We have analysed the proposed circuits by Monte Carlo single-electron circuit simulation, where parasitic capacitance between ground and quantum dots, C_g, and that between the second-neighbor quantum dots, C_d are considered. We found that circuits A and B have almost the same characteristics. Because the structure of circuit-A is simpler than that of circuit-B, we only describe here the simulation results for circuit-A.

The operation temperature ranges for feasible capacitance values are shown in Fig. 4A. The upper limit of operation temperature gradually lowers with increasing C_g and C_d. In order to operate the circuit at higher temperature, one has to scale down all the values of capacitance, and at the same time, scale up the applied voltages. For room-temperature operations, a tunnel junction capacitance of 0.01 aF is required as shown in Fig. 4B although this value is very difficult to realize.

Setting margin of C_o is shown in Fig. 4C as a function of the word length (the number of the connected BCs). There exist minimum values of C_o for the correct operation, and the values increase as the word length increases. This is because some electrons e_R cannot tunnel to node N_o because of the Coulomb-blockade effect by other e_R's. If the parasitic capacitance is negligible, there do not exist the upper limits of C_o. However, C_o should be as small as possible because the sensitivity to one electron in the output voltage e/C_o decreases with increasing C_o.

5. Conclusion

The proposed quantum-dot structures can be realized in the near future because they are similar to floating-gate quantum-dot memory devices that have already been fabricated. However, well-controlled self-organization technologies should be developed.
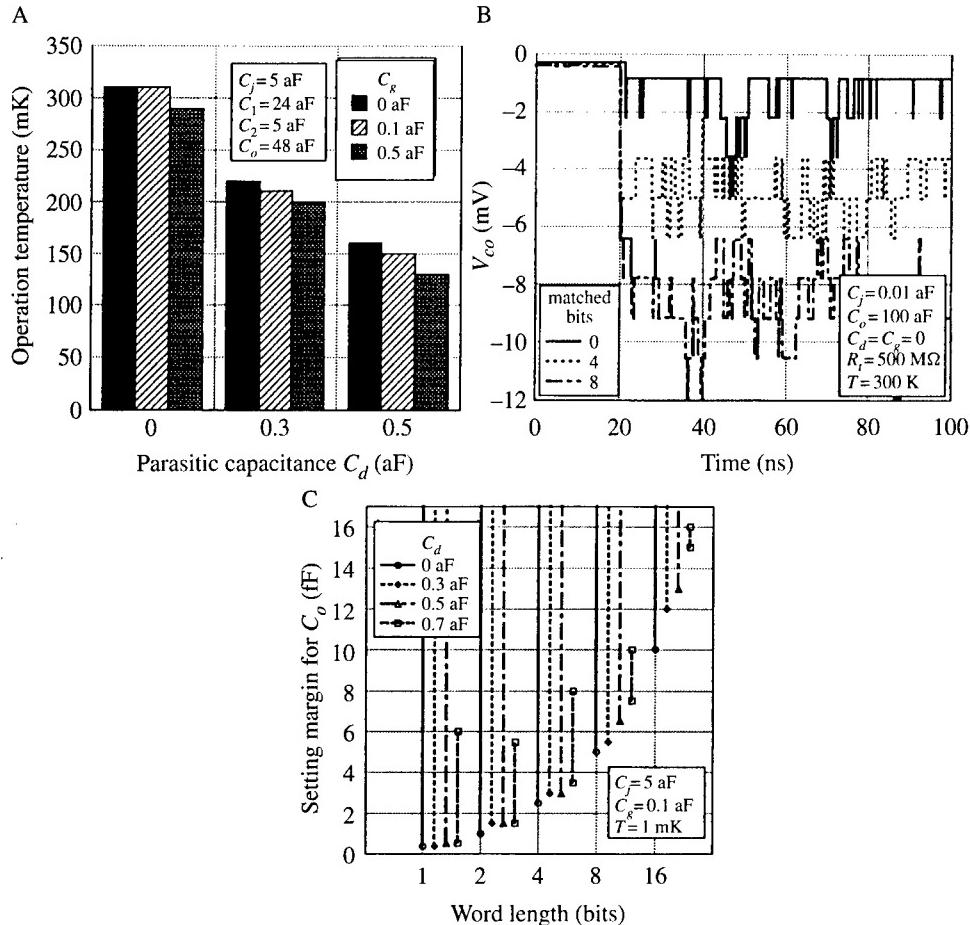
Fig. 4. Simulation results for circuit-A: A, operation temperature range with one BC; B, output voltage changes at room temperature; and C, setting margin for C_o.

References

[1] S. Shimano, K. Masu, and K. Tsubouchi, Jpn. J. Appl. Phys. **38**, 403 (1999).

[2] R. H. Chen, A. N. Korotkov, and K. K. Likharev, Appl. Phys. Lett. **68**, 1954 (1996).

[3] M. Saen, T. Morie, M. Nagata, and A. Iwata, IEICE Trans. Electron. **E81-C**, 30 (1998).

[4] T. Yamanaka, T. Morie, M. Nagata, and A. Iwata, Ext. Abs. of Int. Conference on Solid State Devices and Materials 190 (Japan Society of Applied Physics, Hiroshima, Japan, 1998).

Continuum limit of single-electron tunneling

M. Stopa[†]

Walter Schottky Institute, 1 Am Coulombwall, D-85748 Garching, Germany

We consider the continuum limit of the standard model for treating single-electron tunneling (SET) of electrons through a one-dimensional array of tunnel junctions. We show that the formalism reduces to the computation of the motion of overdamped particles undergoing potential gradient flow, with the potential being given by the full interacting free energy of the electrons in the system. We show that the tunneling coefficients in the SET model can be re-interpreted in terms of a diffusion coefficient and a temperature and that therefore the SET problem reduces to a fully self-consistent treatment of overdamped particle diffusion.

© 2000 Academic Press

Key words: single electron, tunnel junctions, Coulomb blockade.

Single-charge phenomena can be regarded, broadly speaking, as that behaviour which results when the electron *cannot* be adequately described as a *test charge*. A standard formalism has been developed [1] for treating electron transport in certain mesoscopic proto-devices which incorporates the single-charging effects via a parametrization of the energy in terms of capacitances. We will call this the single-electron tunneling (SET) formalism. The theory describes, in particular, electron tunneling to grains and/or through junctions of very small capacitance, i.e. such that $e^2/C > k_B T$.

With regard to charging effects, the salient feature of tunnel junctions, or, more precisely, the islands between tunnel junctions, is their ability to localize the electron on some length scale of order C. This makes the electron, in some sense, classical by providing it with a well-defined position at all times. Indeed, as we will demonstrate here, the SET formalism can be regarded as a discretization of the equations of fully self-consistent transport of overdamped, classical (i.e. localized) particles in a dielectric medium. The demonstration proceeds by examining, in the 'continuum limit' (see below), the two main elements of SET formalism: the electrostatic equation (essentially Poisson's equation) and the equation of motion which, in the SET model, is merely an expression for tunnel rates between neighboring sites.

The system that we study is depicted in Fig. 1. This is a well-studied array of tunnel junctions [2] in one dimension connecting a source and drain, where each island between two junctions has an assumed capacitance to ground C_g and where the tunnel junctions themselves have capacitance C. The electrostatic state of the system is determined by specifying the source and drain potentials, V_s and V_d resp., and the charges on the islands, $q_i = en_i + q_{i0}$ where the mobile charges are integer multiples of e and where the

---

[†]E-mail: stopa@wsi.tum.de
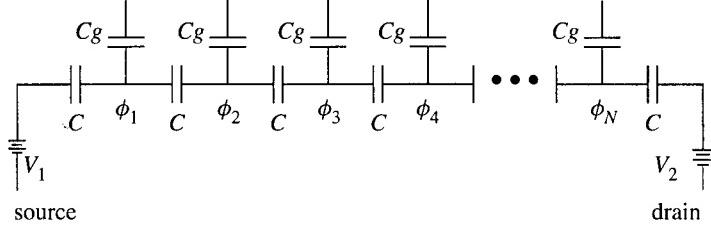
Fig. 1. Schematic of tunneling array described in the text.

background charges q_{i0} are continuous. Kirchoff's laws for the array (Fig. 1) can be applied to immediately obtain as expression for the island potentials ϕ_i in terms of the charges [2],

$$C[2\phi_i - \phi_{i+1} - \phi_{i-1}] + C_g\phi_i = q_i, \tag{1}$$

where $\phi_0 \equiv V_s$ and $\phi_{N+1} \equiv V_d$. As discussed by Bakhvalov *et al.* [2], charges in the arrays are dressed by polarization clouds, the combinations being called 'solitons', with a width $\lambda = \sqrt{C/C_g}$. We introduce a physical length to the array $L \equiv N \Delta x$ and take the limit of eqn (1) as $\Delta x \to 0$ and $N \to \infty$. We demand that L and the soliton length as a fraction of L remain constant. We also take $\rho(x) \equiv q_i/\Delta x$ constant. There is still freedom in taking the limit, which we fix by demanding that $\epsilon \equiv C \Delta x$, the dielectric constant, also remain finite in the limit. The resulting equation is

$$-\frac{\partial^2\phi(x)}{\partial x^2} + k^2\phi(x) = \frac{\rho(x)}{\epsilon}, \tag{2}$$

where the inverse screening length is $k \equiv \sqrt{C_g/\Delta x^2 C}$.

The rate at which electrons tunnel from island i to neighboring island $i + 1$ is expressed, in the low-temperature limit, as [1]

$$\Gamma_{i,i+1} = -\frac{1}{e^2 R_T}\Delta E \; \Theta(-\Delta E), \tag{3}$$

where ΔE is the change in the total energy associated with the tunneling event. Here R_T is the junction resistance which incorporates the barrier tunneling matrix elements as well as the densities of states of the islands. Note that the tunneling rate is zero if the energy change is positive. In the continuum limit we perform a canonical transformation from the island occupancies n_i to the soliton *positions* x_j. The tunneling rates then become, in the limit, soliton *velocities*

$$v_j = \frac{1}{e^2\mathcal{R}}\nabla_x E(\{x_j\}), \tag{4}$$

where we have scaled the velocity with the step length and where $\mathcal{R} \equiv R_T/\Delta x$ is held constant as the limit is taken.

If we ignore the contribution from the background charge for simplicity, the total free energy ('free' because the work from the power supplies is included) is

$$E(\{x_j\}) = \frac{1}{2}\sum_{j=1}^{M} e\phi(x_j) - \frac{1}{2}(Q_s V_s + Q_d V_d) - m_s e V_s - m_d e V_d, \tag{5}$$

where Q_s and Q_d are the charges on the source and drain leads and m_s and m_d are the number of electrons which have (by definition) entered the array from the source and drain, respectively. There are, by assumption, a total of M solitons, but this number obviously changes as charges enter and leave the array.
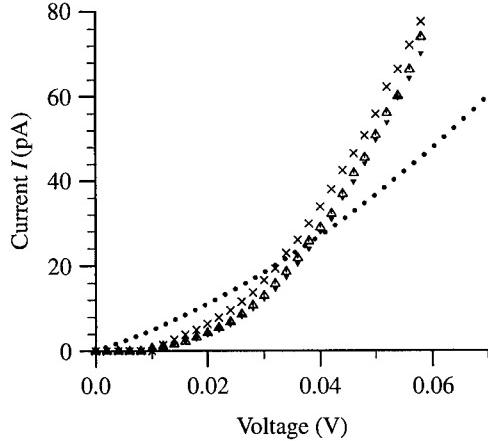
Fig. 2. *I–V* for SET arrays with $N = 20$ (cross), 60 (hollow triangle) and 120 (filled triangle), capacitances scaled as described in the text. Dots indicate drift-diffusion calculations.

Equations (2), (4) and (5) are solved self-consistently and represent the motion of a collection of over-damped, interacting particles undergoing gradient flow. In contrast to the normal case, however, E is the total, interacting free energy of the system, rather than an externally imposed potential.

We have only used the assumption that the particle motion proceeds via tunneling in eqn (4) where we find that the velocity is proportional to the energy gradient. Thus, we may *re-interpret* the proportionality constant as a reciprocal mass-viscosity, $1/m\gamma$ or, via Einstein's relation, we may write

$$v_j = \frac{D}{k_B T} \nabla_{\mathrm{x}} E(\{x_j\}), \tag{6}$$

where D is the diffusion constant of the particles. In summary, we have shown that the SET model calculation of electrons propagating through an array of tunnel junctions is equivalent to a discrete approximation to the motion of overdamped, interacting diffusive particles.

In Fig. 2 we show the current voltage $(I–V)$ characteristics for sample arrays consisting of $N = 20$, 60 and 120 islands. The capacitances and resistances are scaled as described above. Clearly this results in the characteristics for the three cases being essentially the same, although for $N = 20$ there is still slight deviation due to the discreteness of the system. We also exhibit in Fig. 2 the $I–V$ characteristic for the motion of a continuous charged fluid through a constriction as computed through a self-consistent drift-diffusion formalism [3]. An approximate mapping has been employed between the parameters of the SET model (diffusion constant, capacitances) and those of the drift-diffusion model (diffusion constant, mobility, dielectric constant). While the order of magnitude of the resulting currents are approximately the same, there is some residual difference in the shape of the characteristics. In the future we wish to make the mapping between drift-diffusion and SET more rigorous. Since SET is distinct in that it considers discrete charges, rather than a continuum fluid, the numerical disparity between the two models is expected to directly reflect the single-charge behaviour of the system in question.

In conclusion, we have examined the scaling properties of the electrostatic and kinematic equations of the model for single-electron tunneling in arrays of tunnel junctions. We have shown that in the continuum limit the equations reduce to a description of a collection of particles moving interactively in a viscous medium. The potential gradient to which they respond is the total interacting free energy which, for one dimension, can be evaluated very easily when the full Poisson equation is replaced by a capacitance matrix, as is standard in SET.

References

[1] For a general review see H. Grabert and M. H. Devoret, *Single Charge Tunneling* (Plenum, New York, 1992).

[2] N. S. Bakhvalov, G. S. Kazacha, K. K. Likharev, and S. I. Serdyukova, Sov. Phys. JETP **68**, 581 (1989).

[3] M. Stopa, future publication.

Effect of electromagnetic environment effect on coherent and incoherent Cooper pair tunneling in superconducting C-SET

SHUICHI IWABUCHI[†]

Department of Physics, Faculty of Science, Nara Women's University, Kitauoya-Nishimachi, Nara 630-8506, Japan

KANAMI MICHIGAMI, TOSHIE SANO

Department of Physical Science, Graduate School of Human Culture, Nara Women's University, Kitauoya-Nishimachi, Nara 630-8506, Japan

Cooper pair tunneling in voltage-biased superconducting C-SET structure is discussed with emphasis on the electromagnetic environment effect based on the self-consistent microscopic theory of Coulomb blockade in C-SET. It is shown that coherent Cooper pair tunneling survives only in the low impedance limit where charge fluctuation is large, while incoherent Cooper pair tunneling survives in both low- and high-impedance limits.

© 2000 Academic Press

Key words: Coulomb blockade, Cooper pair tunneling, electromagnetic environment effect.

1. Introduction

Tunneling in ultrasmall Josephson double junction is one of the current exciting topics in this field. Coherent charge states realized by coherent Cooper pair tunneling in S-SET (superconducting C-SET) is considered to be used as q-bits in quantum computing [1]. Physics of this system is normally discussed by two parameters E_J/E_c and $\alpha \equiv R_q/R_{qp}$ (E_J: Josephson coupling energy, $E_c \equiv e^2/2C$: charging energy, R_{qp}: resistance for quasi-particle tunneling and $R_q \equiv 2e^2/h$). However, in order to discuss physics due to Coulomb blockade (CB), consideration of electromagnetic environment effect (EMEE), which can be specified by the parameter $\hbar\omega_j/E_c$ ($\hbar\omega_j$:energy of electromagnetic environmental mode), is inevitably important even in the double junctions [2]. We first propose the self-consistent microscopic theory of S-SET for arbitrary environmental impedance. Based on the model we then discuss the effect of CB on Cooper pair tunneling with emphasis on the electromagnetic environment effect while the effect of α is not explicitly considered in this paper (Fig. 1).

[†]Also at: Department of Physical Science, Graduate School of Human Culture, Nara Women's University, Kitauoya-Nishimachi, Nara 630-8506, Japan
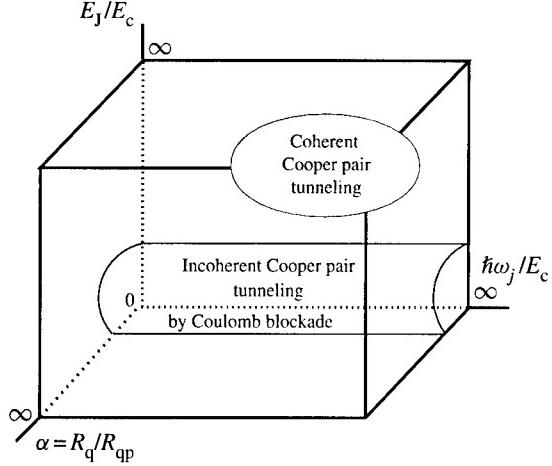
Fig. 1. Competition between Coulomb blockade and Cooper pair tunneling.
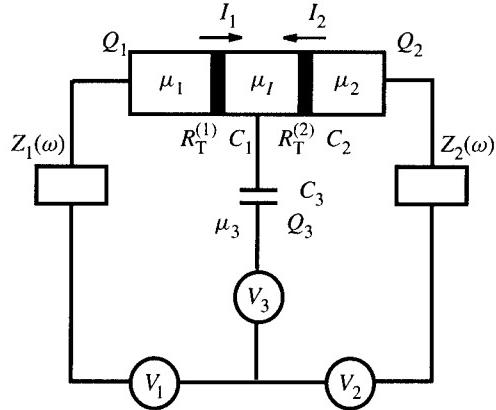
Fig. 2. Voltage-biased S-SET. $R_T^{(i)}$, C_i are tunnel resistance and capacitance of the junction i and Q_i is the charge induced on the electrode i with chemical potential μ_i. $C_\Sigma \equiv \sum_{i=1}^{3} C_i$ and $C \equiv C_1 C_2 / C_\Sigma$. $Z_i(\omega) = i\omega L_i$ etc., for simplicity.

2. Hamiltonian of S-SET

We consider the voltage-biased ultrasmall Josephson double junctions with external circuit as shown in Fig. 2. Hamiltonian of the system is then of form $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_T$. $\mathcal{H}_0$ consists of

$$\mathcal{H}_{el} = \sum_{i=1,2,3,I} \left\{ \sum_{k\sigma} \epsilon_k^{(i)} \, a_{k\sigma}^{(i)\dagger} a_{k\sigma}^{(i)} - \sum_k (\Delta_i^* a_{k\uparrow}^{(i)} a_{-k\downarrow}^{(i)} + \Delta_i a_{-k\downarrow}^{(i)\dagger} a_{k\uparrow}^{(i)\dagger}) \right\} , \tag{1}$$

$$\mathcal{H}_{env} = \sum_{j=1}^{2} \left\{ \frac{(\omega_j/\omega_L)^2}{2L_\Sigma} {\varphi'_j}^2 + \frac{{Q'_j}^2}{2C} - Q'_j V'_{j,1} \right\} , \quad [Q'_i, \varphi'_{i'}] = i\hbar \delta_{i,i'} \quad \text{etc.} \tag{2}$$

$$\mathcal{H}_c = (q/e - n_c)^2, \quad U \equiv e^2/2C_\Sigma , \quad n_c \equiv -\sum_{j=1}^{3} C_j V_j/e + eV_c/(2U) \tag{3}$$

Note that n_c includes the chemical potential of the island and should be determined self-consistently by current continuity condition, $I_1 + I_2 = 0$. Canonical transformation from (Q_i, φ_i) to (Q_i', φ_i') was made to diagonalize $\mathcal{H}_{em}$. Note that Q_3' is related to island charge q as $Q_3' = -\sqrt{C/C_\Sigma} \cdot q$. Therefore Q_3' is a quantized charge since q is quantized as $q|m\rangle = me|m\rangle$ (m:integer) in contrast to Q_1' and Q_2' which have continuous eigenvalues. Correlation functions of φ_i' ($i = 1, 2$) and φ_3' describe EMEE and charged states of the island, respectively. Note that as the increase in degrees of freedom gives rise to more than one environmental mode which leads to the modification of EMEE. $V_{i,I}'$ is also the transformation of $V_{i,I} = -(\mu_i - \mu_I)/e$. Tunneling Hamiltonian is then expressed as

$$
\mathcal{H}_T = \sum_{i=1}^{2} \left\{ \sum_{kk'\sigma} T_{kk'}^{(i)} e^{ie/\hbar \sum_{j=1}^{3} \sqrt{\kappa_i} \eta_{ij} \varphi_i'} a_{k\sigma}^{(i)\dagger} a_{k'\sigma}^{(I)} + h.c. \right\} ,
\tag{4}
$$

where $\kappa_i = C/C_i$ and η is the transformation matrix for diagonalization [3].

3. Cooper pair tunneling

3.1. Coherent Cooper pair tunneling

For $E_J/E_c \gg 1$, coherent Cooper pair tunneling caused by the ordinary Josephson tunneling processes is expected. In this regime CB is not so effective since quantum fluctuation in number is large, so we should not use $|m\rangle$ for evaluating correlation function of φ_3'. Instead, the phase fixed state $|\varphi_3'\rangle \equiv \sum_m c_m |m\rangle$ (c_m:normalization constant) should be used. Since tunneling processes are proportional to $T_{kk'}^{(i)} T_{-k-k'}^{(i)}$ or $T_{kk'}^{(i)*} T_{-k-k'}^{(i)*}$, the current through junction $i (= 1, 2)$, $I_{CCPT}^{(i)}$ then contains the following phase correlation functions: $\mathcal{F}_{i,\pm}^{(env)} \equiv \langle \exp\{\pm ie/\hbar \sum_{j=1,2} \sqrt{\kappa_i} \eta_{ij} \varphi_j'(t)\} \exp\{\pm ie/\hbar \sum_{j=1,2} \sqrt{\kappa_i} \eta_{ij} \varphi_j'(t')\} \rangle_{\mathcal{H}_{env}}$ and $\mathcal{F}_{i,\pm}^{(I)} \equiv \langle \varphi_3'| \exp\{\pm ie/\hbar \sqrt{\kappa_i} \eta_{i3} \varphi_3'(t)\} \exp\{\pm ie/\hbar \sqrt{\kappa_i} \eta_{i3} \varphi_3'(t')\} |\varphi_3'\rangle$. In the high-impedance limit (HIL: $E_c/\hbar\omega_j \gg 1$) and intermediate-impedance limit (IIL: $\hbar\omega_1 \gg E_c \gg \hbar\omega_2$), $I_{CCPT}^{(i)} = 0$, since $\mathcal{F}_{i,\pm}^{(env)} = 0$. It is natural, since phase fluctuation is very large in HIL and IIL. In the low-impedance limit (LIL: $E_c/\hbar\omega_j \ll 1$), on the other hand, $I_{CCPT}^{(i)}$ survives and is reduced to $(2e/\hbar) E_J \sin\phi_i$ ($\phi_i \equiv \theta_i - \theta_I$, $\Delta_i = |\Delta_i| e^{i\theta_i}$) at $\mu_i - \mu_I = 0$ (zero bias on the ith junction) for $U = 0$ (no CB) and $|\Delta_i| = |\Delta_I|$. This is a direct consequence of the fact that the number fluctuation at outer electrodes is very large in LIL.

3.2. Incoherent Cooper pair tunneling

3.2.1. Lowest-order tunneling

Let us consider the regime $E_J/E_c \ll 1$. In this regime the quantum fluctuation of the charge is very small, so the charged state of the island can be described by $|m\rangle$. Therefore, the Josephson current proportional to E_J is strictly forbidden since $\langle m | e^{\pm ie/\hbar \sqrt{\kappa_i} \eta_{i3} \varphi_3'} e^{\pm ie/\hbar \sqrt{\kappa_i} \eta_{i3} \varphi_3'} | m \rangle = 0$ due to the number conservation.

3.2.2. Higher-order tunnelings

Obviously the higher-order Cooper pair tunneling processes must be considered [5]. The lowest non-vanishing contribution is of form

$$
I_{ICCPT}^{(i)} = 2\frac{e}{\hbar^4} \Re_e \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \int_{-\infty}^{t_3} dt_2 dt_3 dt_4
$$
$$
\times \text{Tr}\{ e^{-\beta\mathcal{H}_0} ([[[\mathcal{H}_T^{(i)}(t_1), \mathcal{H}_T^{(i)}(t_2)], \mathcal{H}^{(i)\dagger}(t_3)], \mathcal{H}_T^{(i)\dagger}(t_4)]
$$

$$+[[[\mathcal{H}_T^{(i)}(t_1), \mathcal{H}_T^{(i)\dagger}(t_2)], \mathcal{H}^{(i)}(t_3)], \mathcal{H}_T^{(i)\dagger}(t_4)]$$

$$+[[[\mathcal{H}_T^{(i)}(t_1), \mathcal{H}_T^{(i)\dagger}(t_2)], \mathcal{H}^{(i)\dagger}(t_3)], \mathcal{H}_T^{(i)}(t_4)]]\}/\text{Tr}\{e^{-\beta\mathcal{H}_0}\} . \tag{5}$$

Typical contribution from eqn (5) is written down, in LIL and HIL, as

$$\frac{e}{\hbar}\left(\frac{R_q}{4\pi^2}\right)^2 \sum_{m=-\infty}^{\infty} \frac{e^{\beta[\Omega - U(m-\delta n_c)^2]}}{[E_g + A_m^{(i)}(n_c)]} \left\{ \left(\prod_{\ell=i,\mathrm{I}} \frac{|\Delta_\ell|^2}{R_T^{(\ell)}} \int_{|\Delta_\ell|}^{\infty} \frac{dE_\ell}{\sqrt{E_\ell^2 - |\Delta_\ell|^2}}\right) \right.$$

$$\left. \times \sum_{\nu=\pm 1} \frac{[\theta(\nu_1)f(E_i) - \theta(-\nu_1)(1 - f(E_i))][\theta(\nu_2)f(E_I) - \theta(-\nu_2)(1 - f(E_I))]}{A_m^{(i)}(n_c) - (\nu_1 E_i + \nu_2 E_I)} \right\}^2 , \tag{6}$$

where $f(E_\ell)$ is the Fermi distribution function for quasiparticle with energy E in electrode ℓ, $\theta(x)$ step function, $A_m^{(i)}(n_c) \equiv E_g^{(i)} + 2U(m - \delta n_c) + \mu_I - \mu_i$, $\delta n_c \equiv n_c - \lfloor n_c + 1/2 \rfloor$ ($\lfloor x \rfloor$: the greatest integer less than or equal to x) and $E_g^{(i)} = U$ in LIL or $E_c \sum_{j=1,2} \kappa_i \eta_{ij}^2 + U$ in HIL, respectively. Note that there is neither phase dependence nor time dependence. This is due to the fact that tunneling processes such as $T_{kk'}^{(i)} T_{-k-k'}^{(i)} T_{k_1 k_1'}^{(i)*} T_{-k_1-k_1'}^{(i)*}$ only give rise to matrix elements proportional to the time difference and $|\Delta_i \Delta_I|^2$. Therefore, it is not adequate to start from Hamiltonians of form $E_J \cos\phi + \mathcal{H}_T$ [4], since the lowest-order tunneling which depends on ϕ does not exist. Dominant contribution is given at resonant condition $A_m^{(i)}(n_c) = 0$, which can be satisfied with certain sets of ($V \equiv V_1 - V_2$, V_3) for given m. For $\Delta_i = \Delta_I$, at low temperatures, $I_{\text{ICCPT}}^{(i)} \sim (2e/\hbar) E_J \tanh^2(\beta|\Delta|/2)/E_g^{(i)}$ for (V, V_3) satisfying $|\mu_i - \mu_I|/e \sim E_g^{(i)}/e$. The present theory provides a more fundamental starting point to discuss the Cooper pair tunneling of this system.

4. Summary and discussions

We proposed a theory which can treat both coherent and incoherent Cooper pair tunnelings in S-SET for arbitrary environmental impedance. It was shown that coherent Cooper pair tunneling exists only at the low-impedance limit, while incoherent Cooper pair tunneling survives in the entire region from high- to low-impedance limits. It should be noted that, to discuss the latter, a microscopic Hamiltonian is needed since the lowest-order Cooper pair tunneling is strictly forbidden. To incorporate the effect of energy dissipation explicitly, quasi-particle tunneling should be also taken into account. Furthermore, to discuss the current–voltage characteristics of the system, the current continuity conditions are explicitly considered for various tunneling schemes such as successive Cooper pair tunnelings and Josephson-quasiparticle cycle tunneling etc. Studies along such directions will be reported elsewhere.

References

[1] Y. Nakamura, C. D. Chen, and J. S. Tsai, Phys. Rev. **79**, 2328 (1997).
[2] H. Higurashi, S. Iwabuchi, and Y. Nagaoka, Phys. Rev. **B51**, 2387 (1995).
[3] S. Iwabuchi, H. Higurashi, F. Wakaya, and Y. Nagaoka, (submitted)
[4] D. V. Averin and V. Ya. Aleshkin, JETP Lett. **50**, 367 (1989).
[5] Shuichi Iwabuchi, *Quantum Coherence and Decoherence*, edited by K. Fujikawa and Y. A. Ono (Elsevier Science, 1996) p. 139.

NOTICE

Under the terms of the U.S. copyright law effective January 1, 1978, and the copyright laws of many other countries, journal publishers are required to obtain written confirmation from authors to acquire copyright rights for papers published in their journals. We ask your cooperation because it is of vital importance for the widest possible dissemination of your paper. We must have your written confirmation to authorize libraries and other information centers to use this material and to perform other appropriate publishing functions. We are therefore asking you to sign and return the applicable section of the form enclosed with this notice. PLEASE SIGN ONLY THE APPLICABLE SECTION.

Which section of the form should you sign? If you and your co-authors, if any, own rights in your paper as individuals (for example, you are working at a university or research institute, even if you are supported by a government grant), please sign **Section A** of the attached form. If you work for a corporation that owns the rights in the paper, please have an authorized officer of the corporation sign **Section B**. If the paper is in the public domain because you and all your co-authors created it within the scope of your employment with the U.S. government, please sign **Section C**. If the paper is subject to Crown copyright because you and all your co-authors are employees of the British or Canadian government, please sign **Section C**. If not all authors are government employees, then the non-government author should sign **Section A**. If any governmental entity other than the U.S., British or Canadian government owns copyright, please call Academic Press at the number below. No manuscript will be processed until the appropriate form has been returned.

If your manuscript is not published within a reasonable period of time after acceptance, upon written request, Academic Press will promptly return the rights to the manuscript to the author(s).

General Terms of Publication in Academic Press Journals

1. *Author Warranties*

1.1. The work submitted for review is new and has been written by the stated authors and has not been published elsewhere.

1.2. The work submitted is not currently under review for another journal and will not be submitted to another journal while under review for the present journal.

1.3. Submission of the work has been approved by all the authors and by the institutions where the work was carried out. Written authorization may be required at the Editor's discretion. For works with multiple authors, all the authors participated meaningfully in the work and agree on the submitted version of the work.

1.4. Any person cited in the work as a source of personal communications has approved such citation.

1.5. Submitted paper and electronic versions of the work are identical, except for elements that cannot be represented on paper.

1.6. All necessary permissions for the reproduction or excerpting of the work of others have been obtained.

2. *Understandings*

2.1. All works submitted to Academic Press journals are subject to peer review, a process that is the responsibility of the Editors.

2.2. The authors agree that, if the work is accepted for publication, they will transfer the copyright in it exclusively to Academic Press, including the right of reproduction in all forms and media, whether now known or hereafter developed, and the right to include it in collections and databases.

3. *Personal Servers*

3.1. Upon submitting an article to an Academic Press journal for review and possible publication, the authors are requested to add the following notice to the first screen of any posted electronic preprint versions of the paper: *This work has been submitted to Academic Press for possible publication. Copyright may be transferred without notice, after which this version may no longer be accessible.* Authors should note that posting the entire work may be regarded as prior publication by some journal editors (see Information for Authors of the specific journal).

3.2. When an Academic Press journal accepts the work for publication, the authors may post it, in its final accepted form, on their personal servers (but not on any organized preprint server) with a notice *Accepted for publication in* ⟨name of journal⟩ *as of* ⟨date⟩, until it is published by Academic Press in print or electronic form.

3.3. After publication, authors may post their Academic Press copyrighted material on their own servers without permission, provided that the server displays as the first line of the HTML page the following notice alerting readers to their obligations with respect to copyrighted material: *This material has been published in* ⟨name of journal, issue number and date, page number⟩, *the only definitive repository of the content that has been certified and accepted after peer review. Copyright and all rights therein are retained by Academic Press. This material may not be copied or reposted without explicit permission.* The posted work must also include the Academic Press copyright notice (*Copyright © 200x by Academic Press*) and a link to IDEAL (International Digital Electronic Access Library) at ⟨http://www.idealibrary.com⟩.

IMPORTANT: Note that Academic Press may choose to publish an abstract or portions of the paper prior to publishing the paper in the journal. Please contact Academic Press immediately if you would like Academic Press to refrain from such prior publication for any reason, including disclosure of a patentable invention.

If you have any questions about the procedures, please call Cindy MacDonald at Academic Press, Tel.: +1(619) 699-6413.

Please **return the form** to *Superlattices and Microstructures*, Academic Press, 32 Jamestown Road, London NW1 7BY, UK.

Copyright Transfer Agreement

Journal: SUPERLATTICES & MICROSTRUCTURES

Tentative Manuscript Title: _____

_____

Authors: _____

A. University/research institute employees and others who own rights as individuals. You hereby confirm the assignment to Academic Press (AP) of the entire copyright in and to the manuscript named above (the Work) throughout the world in all forms and media, whether now known or hereafter developed, and in all languages effective if and when it is accepted for publication by AP. You agree to abide by the terms of publication stated on the accompanying notice. You also warrant that the Work has not been published before and contains no materials the publication of which would violate any copyright or other personal or proprietary right any any person or entity, and you acknowledge that AP is relying on this document in publishing this Work. Notwithstanding the assignment of copyright to AP, you:

1. Retain patent rights, if any.
2. May make photocopies of all or part of the Work for use in your classroom teaching.
3. May use your original figures/tables/illustration/photographs from the Work in future works of your own.
4. May include the Work as part of your dissertation, for non-commercial distribution only.

You agree that all copies and inclusions made under the above conditions will carry a notice of copyright and a citation to the journal article.

To be signed by all the authors (or by an author who has obtained the written assent of the non-signing authors).

Signed: _____ Date: _____

B. Corporate employees. The Work was written as part of the duties of the author as an employee of the comapny signing below or otherwise as a "work made for hire" (as defined in the relevant copyright law) for such company. As the authorized representative of such company, you hereby confirm the assignment to AP of the entire copyright in and to the Work throughout the world in all forms and media, whether now known or hereafter developed, and in all languages effective if and when it is accepted for publication by AP. You agree to abide by the terms of publication stated on the accompanying notice. You also warrant that the Work has not been published before and contains no materials the publication of which would violate any copyright or other personal or proprietary right of any person or entity, and you acknowledge that AP is relying on this document in publishing this Work. Notwithstanding the assignment of copyright to AP, you:

1. Retain patent rights, if any.
2. May prepare derivative works based on the Work, with proper acknowledgment.
3. May reproduce the Work in reasonable quantities for internal use, with proper acknowledgment. However, copies of the Work may not be used for resale or in any way that implies that AP, the journal, or the editorial board endorses any procedures or products of your company.

Signed: _____ Date: _____

Title: _____

C. Government employees. You certify that the Work has been written as part of your official duties either as (a) an employee of the U.S. government or another government whose works are not subject to copyright protection, and as such there is no copyright to transfer; or as (b) an employee of the British, Canadian, or other government whose works are subject to Crown copyright, and via the authorized signature of the relevant government entity below, AP is hereby obtaining a non-exclusive license to publish (and to authorize others to publish) the Work in all forms and media, whether now known or hereafter developed, and in all languages effective if and when it is accepted for publication. If not all of the authors are government employees, one of the nongovernment employees should sign Section A. You also warrant that the Work has not been published before and contains no materials the publication of which would violate any copyright or other personal or proprietary right of any person or entity, and you acknowledge that AP is relying on this document in publishing this Work.

To be signed by all the authors (or by an author who has obtained the written assent of the non-signing authors).

Signed: _____ Date: _____

Government authorization (if needed): _____ Date: _____

Instructions to Authors

Either a manuscript or a printed copy of a compuscript should be submitted to Supriyo Datta, c/o Academic Press, Block A2, Westbrook Centre, Milton Road, Cambridge CB4 2YG, U.K. Please include your electronic mail address which will be used for all communication regarding your manuscript, unless you indicate otherwise. You will receive an acknowledgement upon receipt of your manuscript and can expect a communication from the Editor within four weeks. Inquiries regarding the status of the manuscript can be sent to superlattices@harcourt.com mentioning the name of the first author and the title of the paper. It is recommended that you include with your cover letter a brief description (less than 200 words) of how this paper advances the state of knowledge in its field along with a list of five reviewers (with contact information if possible) who are qualified to judge your manuscript. This will greatly aid the editor in evaluating your manuscript, although the editor may choose a reviewer from outside your list.

Preparation of compuscripts. If you decide to submit a compuscript in LATEX, please prepare your manuscript according to the guidelines that you will find on the homepage. To access this site go to http://www.academicpress.com/sm. Here you will find the LaTeX template and Style Files in the information box. In case of problems, please contact superlattices@harcourt.com. Note that the compuscript itself is not submitted for refereeing to the editor, only a hard copy. Instructions for submissions of compuscript are given when the paper is accepted for publication by the editor.

Preparation of typed manuscripts. If you decide to submit a typed manuscript, papers should be typed in double spacing throughout, including tables, footnotes, references and legends to tables and figures. One side of the paper, only, should be used and there should be a margin of at least 2.5 cm all round. The position of tables and figures should be clearly indicated, in sequence, in the text. Tables, footnotes and captions to figures should be typed separately. Sections and sub-sections may be numbered, and the decimal system should be used, i.e. 1., 1.2., 1.1.2., etc.

Abstract. An abstract will be printed at the head of all papers: this should not exceed 300 words, and should be intelligible to the general reader without reference to the main text. Abbreviations should be avoided.

Digital object identifier. Academic Press now assigns a unique digital object identifier (DOI) to every article it publishes. The DOI appears on the title page of the article. It is assigned after the article has been accepted for publication and persists throughout the lifetime of the article. Due to its persistence, it can be used to query Academic Press for information on the article during the production process, to find the article on the Internet through various Web sites, including IDEAL, and cite the article in academic references. When using an Academic Press article in a reference section, it is important to include the article's DOI in the reference as volume and page information is not always available for articles published online. The References section, below, shows samples of DOI included in references. Further information can be found at http://www.academicpress.com/doi

References. References in the text are to be indicated by consecutive numbers in square brackets. The reference list at the end of the paper should take the following form. When citing an Academic Press journal, include the digital object identifier (DOI), if noted, from the article's title page. Please note the following examples.

[1] B. Tanatar and A. K. Das, Superlatt. Microstruct. **27**, 23 (2000), doi:10.1006/spmi.1999.0815.
[2] H. Kümmel, R. Till and A. Philipp, Superlatt. Microstruct. doi:10.1006/spmi.1999.0820.
[3] H. A. Blackstead and J. D. Dow, "High-temperature superconductivity," in *Proceedings of the Second International Symposium on Quantum Confinement Physics and Applications*, edited by M. Cahay, S. Bandyopadhyay, J. P. Leburton, A. W. Kleinsasser, and M. A. Osman, (Electrochemical Society, Pennington, NJ, 1994), Vol. 94-17, p. 408, especially p. 410.

Note that the official abbreviations of the journal titles may be used, or else the entire title. Unpublished preprints should be cited by title and journal (if known). Book titles are in italics. Author lists may be shortened to "*et al.*" in the text, but never in the reference list.

Figures. Authors' illustrations should be supplied as electronic files. Figures should be saved in either Encapsulated PostScript (EPS) or TIFF formats. Explanations and keys should be placed in the captions as far as possible. Attached firmly to each figure, on a separate sheet, mark, e.g. "Fig. 2 of 6, Blackstead and <u>Dow</u>, MS 123", where the name of the corresponding author is underlined, and the manuscript number 123 is supplied by the editor upon manuscript acceptance. Colour figures are permitted at a cost of £600 sterling per page, payable by the authors.

Equations. In manuscripts, all variables should be typed in italics, if possible. In compuscripts, use the LATEX math font. Numbering of equations is optional.

Copyright/offprint forms. Authors submitting a manuscript do so on the understanding that if it is accepted for publication, exclusive copyright in the paper shall be assigned to the Publisher. In consideration for the assignment of copyright, the Publisher will supply 25 complimentary offprints of each paper. Further offprints may be ordered at extra cost after acceptance of manuscripts when an order form will be supplied to authors. The Publishers will not put any limitations on the personal freedom of the author to use material contained in the paper in other works.

Papers are accepted in the Journal on the understanding that they have not been or will not be published elsewhere in the same form in any language.

Quantum Structures

Self-Assembled Structures

Generalized Transport Problems

Carbon Nanotubes

Surfaces and Analysis

SETs and Quantum Circuits

Superlattices and Microstructures